Software Engineering Institute

Management and Education of the Risk of Insider Threat (MERIT): Mitigating the Risk of Sabotage to Employers' Information, Systems, or Networks

Dawn M. Cappelli
Akash G. Desai (Information Networking Institute, Carnegie Mellon University)
Andrew P. Moore
Timothy J. Shimeall
Elise A. Weaver (Worcester Polytechnic Institute; CERT Visiting Scientist)
Bradford J. Willke

March 2007

TECHNICAL NOTE
CMU/SEI-2006-TN-041

CERT Program
Unlimited distribution subject to the copyright.

Carnegie Mellon

This work is supported by the Army Research Office through grant number DAAD19-02-1-0389 (“Perpetually Available and Secure Information Systems”) to Carnegie Mellon University's CyLab.

This report was prepared for the

SEI Administrative Agent
ESC/XPK
5 Eglin Street
Hanscom AFB, MA 01731-2100

The ideas and findings in this report should not be construed as an official DoD position. It is published in the interest of scientific and technical information exchange.

This work is sponsored by the U.S. Department of Defense. The Software Engineering Institute is a federally funded research and development center sponsored by the U.S. Department of Defense.

Copyright 2007 Carnegie Mellon University.

NO WARRANTY

THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

Use of any trademarks in this report is not intended in any way to infringe on the rights of the trademark holder.

Internal use. Permission to reproduce this document and to prepare derivative works from this document for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.

External use. Requests for permission to reproduce this document or prepare derivative works of this document for external and commercial use should be addressed to the SEI Licensing Agent.

This work was created in the performance of Federal Government Contract Number FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013.

For information about purchasing paper copies of SEI reports, please visit the publications portion of our Web site (http://www.sei.cmu.edu/publications/pubweb.html).

Table of Contents

Acknowledgements **vii**

Abstract **ix**

1 Introduction **1**

2 Simulation-Based Learning **3**

3 Origins of CERT Research **5**
3.1 Key Findings from the Insider Threat Study 6
3.2 Targeted Lessons for Training 7

4 Fictional Case for Training: iAssemble, Inc. **9**

5 Methodological Background **11**
5.1 Notation 12

6 Model Assumptions **16**
6.1 Scope 16
6.2 Organization 16
6.3 Insiders 16
6.4 Access 17
6.4.1 Access paths 17
6.5 Defenses 18

7 Modeling Behavioral Aspects **20**
7.1 Insider Expectation of Freedom 20
7.2 Escalation of Disgruntlement and Sanctioning 23

8 Modeling Technical Attack Aspects **27**
8.1 Attack Setup and Concealment 27
8.2 Attack Escalation 28

9 Modeling Technical Defense Aspects **31**

10 Exhibiting the iAssemble Reference Mode **34**

11 Conclusion **37**

Appendix A: The iAssemble Training Case **39**

Appendix B: Simulation Model Overview **45**

References **47**

List of Figures

Figure 1: A Simple Feedback Loop 13

Figure 2: (a) Project Management—Desire to Use Overtime to Correct Schedule; (b) Closed-Loop Representation Showing (Balancing) Feedback to Improve Progress 14

Figure 3: Unintended Burnout Due to Overtime 14

Figure 4: Expected Freedom by Insider 21

Figure 5: Expected and Actual Freedom Growth with Lax Supervisor 22

Figure 6: Expected and Actual Freedom with Strict Supervisor Hired at Week 20 23

Figure 7: Escalation of Disgruntlement and Sanctioning 25

Figure 8: Escalation of Disgruntlement and Sanctioning with Minimal Intervention 26

Figure 9: Disgruntlement and Sanctioning with Proactive Intervention 26

Figure 10: Access Path Stocks and Flows 27

Figure 11: Attack Setup and Concealment 28

Figure 12: Attack Escalation 29

Figure 13: Attack Simulation 30

Figure 14: Risk-Based Auditing and Access Path Disabling 32

Figure 15: Attack Simulation with Audit Quality at 50% 33

Figure 16: Attack Simulation with Audit Quality at 80% 33

Figure 17: Exhibiting the iAssemble Problematic Behavior 34

Figure 18: Function Defining Effect of Access Paths on Access Control Quality 35

Figure 19: Explanation for iAssemble Attack (Simulation) 36

Figure 20: Insider's Method of Attack 42

Figure 21: Explanation for iAssemble Attack 42

Figure 22: Analysis of Insider Access Level 43

List of Tables

Table 1: Simulation Effects of Policy Levers 18

Acknowledgements

CERT would like to thank the Army Research Office and Carnegie Mellon University's CyLab for funding this project.

CERT appreciates the work and dedication of the Insider Threat Study team; without the study none of our follow up on insider threat research would have been possible. Many thanks to the Insider Threat Study research staff:

Carnegie Mellon University, Software Engineering Institute, CERT: Andrew Moore, Bill Wilson, Bradford Willke, Casey Dunlevy, Chris Bateman, Dave Iacovetti (USSS/CERT Liaison), David Mundie, Dawn Cappelli, Mark Zajicek, Stephanie Rogers, Tim Shimeall, Tom Longstaff, Wayne Peterson (USSS/CERT Liaison), Cornelius Tate (USSS/CERT Liaison).

U.S. Secret Service, National Threat Assessment Center: Brandi Justice, Diana McCauley, Eileen Kowalski, Georgeann Rooney, Jim McKinney, Lea Bauer, Lisa Eckl, Marisa Reddy Randazzo, Megan Williams, Michelle Keeney, Susan Keverline, Tara Conway.

CERT would like to acknowledge the valuable participation in development of the Management and Education of the Risk of Insider Threat (MERIT) insider IT sabotage model by the U.S. Department of Defense Personnel Security Research Center (PERSEREC) team: Dr. Lynn Fischer (PERSEREC project sponsor), Dr. Katherine Herbig (PERSEREC), Dr. Eric Shaw (Consulting and Clinical Psychology, Ltd.), and Dr. Stephen R. Band. The PERSEREC team members bring unmatched expertise in insider threat, psychology, espionage, and cyber crime.

CERT would like to thank members of the Security Dynamics Network, a collaborating network of institutions and associated researchers using system dynamics modeling to explore risk dynamics, especially with respect to cybersecurity. Its members include University of New York at Albany; Agder University College; TECNUN, University of Navarra; Worcester Polytechnic Institute; Sandia National Labs; Argonne National Labs; and Carnegie Mellon University. Network members have met several times per year since 2004, providing feedback on various system dynamics projects related to cybersecurity.

Last, but not least, we thank the anonymous reviewers of this technical note and Claire Dixon of the Software Engineering Institute for their valuable suggestions for improving it.

Abstract

The Insider Threat Study, conducted by the U.S. Secret Service and Carnegie Mellon University's Software Engineering Institute CERT Program, analyzed insider cyber crimes across U.S. critical infrastructure sectors. The study indicates that management decisions related to organizational and employee performance sometimes yield unintended consequences that increase risk of insider attack. The problem is exacerbated by a lack of tools for understanding insider threat, analyzing risk mitigation alternatives, and communicating results. The goal of Carnegie Mellon University's Management and Education of the Risk of Insider Threat (MERIT) project is to develop such tools. MERIT uses system dynamics to model and analyze insider threats and produce interactive learning environments. These tools can be used by policy makers, security officers, information technology and human resource personnel, and management. The tools help these users to understand the problem and assess risk from insiders based on simulations of policies, and on cultural, technical, and procedural factors. This technical note describes the MERIT insider threat model and simulation results.

1 Introduction

Insiders, by virtue of legitimate access to their organizations' information, systems, and networks, pose a significant risk to employers. Employees experiencing financial problems have found it easy to use the systems they use at work everyday to commit fraud. Other employees, motivated by financial problems, greed, or the wish to impress a new employer, have stolen confidential data, proprietary information, or intellectual property from their employers. Lastly, technical employees, possibly the most dangerous because of their intimate knowledge of their organizations' vulnerabilities, have used their technical ability to sabotage their employers' systems or networks in revenge for negative work-related events.

In January 2002, the Carnegie Mellon University Software Engineering Institute's CERT Program (CERT) and the United States Secret Service (USSS) National Threat Assessment Center (NTAC) started a joint project, the Insider Threat Study.[1] The study combined NTAC's expertise in behavioral psychology with CERT's technical security expertise to provide in-depth analysis of approximately 150 insider incidents that occurred in critical infrastructure sectors between 1996 and 2002. Analysis included perusal of case documentation and interview of personnel involved in the incident.

Two reports have been published to date as part of the Insider Threat Study, one analyzing malicious insider incidents in the banking and finance sector (Randazzo 2004), and one analyzing insider attacks across all critical infrastructure sectors where the insider's intent was to harm the organization, an individual, or the organization's data, information system, or network (Keeney 2005). We expect two additional reports to be published in 2007: one specific to the information technology and telecommunications sector, and one for the government sector.

The reports include statistical findings and implications regarding technical details of the incidents; detection and identification of the insiders; nature of inflicted or intended harm; as well as insider planning, communication, behavior, and characteristics. The reports have been well received across several stakeholder domains including the business community, technical experts, and security officers. Our fear is that practitioners will mistakenly interpret the results as stand-alone statistics and assign consideration of individual implications to various departments within the organization instead of taking a holistic, enterprise-wide approach to mitigating insider threat risk.

The results of the Insider Threat Study show that to detect insider threats as early as possible or to prevent them altogether, members of management, IT, and human

[1] The Insider Threat Study was funded by the USSS, as well as the Department of Homeland Security, Office of Science and Technology, which provided financial support for the study in fiscal years 2003 and 2004.

resources, and security officers and others in the organization must understand the psychological, organizational, and technical aspects of the problem, as well as how to coordinate their actions over time. CERT staff felt strongly that an important next step in our insider threat research was development of innovative communication, education, and training materials to address this issue. After researching potential methods and tools that could be used for this purpose, system dynamics was chosen for its strengths in modeling and simulation of complex problems.

This paper describes the MERIT project. MERIT stands for the Management and Education of the Risk of Insider Threat. The project goal is to develop a system dynamics model that can be used to better communicate the risks of the insider sabotage threat to an organization's information, systems, or networks. This paper presents that model, a system dynamics model of insider IT sabotage. The model is not a predictive model, but a descriptive one. We formulated this model by analyzing the common patterns across the insider IT sabotage cases, and its purpose is to illustrate the primary patterns that recur in those cases. The output of the model simulation is presented in the context of a fictional organization and is not meant to necessarily apply to any particular organization. We structure the paper as follows:

- Section 2 motivates the use of modeling and simulation for learning about complex systems, such as the problem of insider sabotage and its mitigation.
- Section 3 describes in greater detail the Insider Threat Study and related research being conducted by CERT.
- Section 4 describes a fictional case used as a basis for the MERIT training materials and model development; the case is described more fully in Appendix A.
- Section 5 describes the system dynamics methodology used as a basis for our modeling and simulation.
- Section 6 describes the primary assumptions and scope of the MERIT model.
- Sections 7 through 9 describe the behavioral, attack, and defense aspects of the model, respectively. Appendix B contains a comprehensive overview of the MERIT model.
- Section 10 shows how the model can be used to generate the problematic behavior of the fictional case.
- Section 11 describes conclusions that can be drawn from the current model, and work that remains to be done.

2 Simulation-Based Learning

Any organization using a computerized network represents a complex system involving both people and technology. Each employee navigates the network with a unique level of access and set of authorized capabilities that may change over time. Each organization exercises a policy to optimize productivity while guaranteeing security, whether or not such a policy is well developed and implemented. While there are many such systems, and thus a wealth of common experience upon which to base learning the important components of a good security policy, it is not a given that all organizations manage this learning well.

Sterman describes three challenges to implementing good policy based on lessons learned from interactions in complex systems (Sterman 2006). First, one cannot draw lessons learned unless one has good data. Second, one cannot learn from experience, even with good data, unless one can derive good lessons from that data. Finally, one cannot implement good policy based on lessons learned, unless stakeholders in the system are involved politically in policy development.

Sterman makes compelling arguments that all three of these efforts are hindered by the characteristics of complex systems. Unlike the scientific laboratory, in which variables can be isolated and controlled, the networked organization presents serious challenges to the gathering of valid, reliable, and easily interpretable data from which one can draw clear conclusions. Unambiguous data are difficult to gather because a complex system, such as a networked organization, involves cause-and-effect interactions across many time scales, locations, and areas of expertise. In addition, complex systems not only respond to the decisions taken by the learner but to decisions taken by other agents in response. Finally, one cannot afford to make risky extreme decisions even though these might yield good data for learning, because of the ethical and logistical implications of experimenting with real organizations. As a result, the specific results of any action taken are difficult to discern.

Even with good data, learning isn't guaranteed. To derive lessons learned, one would first form hypotheses about system behavior, gather data, and reflect on whether the results matched those predicted. Finally, one would update one's mental model after reflecting thoroughly on any discrepancies, in a form of double-loop learning (Argyris 1974).

Sterman reviews research from the psychological literature, for example from Wason and from Kahneman and Tversky, indicating that people are not normally prone to gather data that might disconfirm their hypotheses (Wason 1966; Kahneman 1982), especially not in public (Tedeschi 1981) or when working in groups (Janis 1977). When making judgments, we tend to use simplifying rules of thumb that may be efficient but are often biased (Kahneman 1982). Finally, not only do we avoid looking where disconfirming evidence might be, we also respond in ways

that force the system to exhibit types of behavior that confirm our earlier biased beliefs (Rosenthal 1968).

Sterman advocates the use of virtual worlds to overcome our inability to learn in complex systems (Sterman 2006). These need not be computer simulations, as they can be role-play games or physical model environments as well. In simulations, distinguished from games by a verisimilitude that enables knowledge transfer beyond the environment (Lane 1995), it is possible to set up a closed environment with known assumptions. One can then test the impact of policies without the distortion of statistical error and reflect on the outcomes resulting only from one's own actions.

Lane reviews the history of such simulations and describes the value of using these to provide managers with an intellectually and emotionally rich and engaging educational experience (Lane 1995). Lane points that the low cost and unambiguous feedback afforded can be "more helpful than reality" as long as certain caveats are considered. The model should represent the relevant environment with fidelity, the simulation instructions should be clear, the simulation objective (such as maximizing productivity and minimizing risk and cost) measurable and known to the user, and there must be an opportunity for debriefing or reflection. In addition, Lane notes that simulations provide common metaphors for communication about insights or lessons learned.

Groessler elaborates on 15 issues that should be considered when designing such simulations for training (Groessler 2004).

- The first five concern the characteristics of the model, and pertain to how well it balances the fidelity to the context with the necessity for simplification so that lessons can be learned. The model should be validated against real cases without so much complexity as to overwhelm the user.
- The second five address the characteristics of the trainees, balancing the cognitive complexity of the task with the users' learning styles. The simulation should be part of a larger interactive learning environment allowing individuals many ways to glean insights from using it.
- The third five address whether the interactive learning environment encourages good engagement with the task and reflection on lessons learned. Users of the simulation should have the opportunity to monitor indicators of success and should be given opportunities to reflect on their hypotheses and the results of their experiments.

3 Origins of CERT Research

The Security Dynamics Network is a collaborating network of institutions and associated researchers using system dynamics modeling to explore risk dynamics, with a focus on cybersecurity. Its members include the University of New York at Albany; Agder University College; TECNUN, University of Navarra; Worcester Polytechnic Institute; Sandia National Labs; Argonne National Labs; and Carnegie Mellon University. The network was created in 2004 after members of the group convened several workshops to model various aspects of the insider threat (Melara 2003; Anderson 2004; Rich 2005).

Convinced that system dynamics modeling was a viable mechanism for transitioning our insider threat knowledge, CERT sought funding for development of a prototype interactive learning environment (ILE) based on empirically validated models of the insider threat problem developed using Insider Threat Study data.[2] The purpose of MERIT is to develop an ILE using system dynamics for hands-on analysis of the effects of policy, technical, and countermeasure decisions on malicious insider activity. It will provide a means to communicate insider threat risks and tradeoffs, benefiting technical and non-technical personnel, from system administrators to corporate CEOs. The MERIT project was funded by CyLab at Carnegie Mellon University.

At about the same time the MERIT project was initiated, the CERT insider threat team was funded by the U.S. Department of Defense Personnel Security Research Center (PERSEREC) for another system dynamics modeling project. That work is part of an ongoing partnership between CERT and PERSEREC in response to recommendations in the 2000 Department of Defense (DoD) Insider Threat Mitigation report (https://dssacdsws.dss.mil/is201docs/DoD_Insider_Threat_Mitigation.pdf). The purpose of the PERSEREC/CERT project is to develop two system dynamics models based on actual case data—one for insider IT sabotage and one for espionage—and then compare and contrast the models. The comparison could identify countermeasures that could be useful for mitigating both risk of insider IT sabotage and espionage in the DoD. Initial results from this work are published in a technical report by Band and associates [Band 2006].

CERT researchers believed that the scope of MERIT should initially be limited to a well-defined subset of the 150 cases from the Insider Threat Study. Because a model of insider IT sabotage could be used for both the MERIT and PERSEREC projects, the CERT research team decided to first focus MERIT on insider IT sabotage cases. As a result, a base model is being developed for insider IT sabotage that can be used for both projects.

[2] An *interactive learning environment* (ILE) is a process for educational learning that allows the instructor and student to negotiate the context of the curriculum in real time.

One unique aspect of the Insider Threat Study that was a key to its success was the equal attention given to both the technical and psychological aspects of the problem. MERIT enabled the CERT team to realize unexpected benefits from the overlap with the PERSEREC project. CERT's technical security expertise was augmented with expertise from several organizations in the areas of psychology, insider threat, espionage, and cyber crime. Therefore, the system dynamics model for insider IT sabotage being developed for both MERIT and PERSEREC benefits from a broad range of experience regarding the technical, psychological, and organizational factors influencing insider threat risk.

3.1 KEY FINDINGS FROM THE INSIDER THREAT STUDY

We base our system dynamics models on findings from the Insider Threat Study, in particular those cases involving insider sabotage. These were among the more technically sophisticated attacks perpetrated in the study and resulted in substantial harm to organizations. In the 49 cases studied, 81% of the organizations that were attacked experienced a negative financial impact as a result of insider activities. The losses ranged from a low of $500 to a high of "tens of millions of dollars." Seventy-five percent of the organizations experienced some impact on their business operations. Twenty-eight percent of the organizations experienced a negative impact to their reputations. The statistics in this section come from the CERT/USSS Insider Threat Study report on insider sabotage (Keeney 2005).

The first step taken in modeling insider IT sabotage was to identify the key findings to be reflected in our system dynamics model. Below is a summary of the findings:

Insiders were disgruntled and motivated by revenge for a negative work-related event. Fifty-seven percent of the insiders who committed IT sabotage were disgruntled. Eighty-four percent were motivated by revenge, and 92% of all of the insiders attacked following a negative work-related event such as termination, dispute with a current or former employer, demotion, or transfer.

Insiders exhibited concerning behavior prior to the attack. Eighty percent of the insiders exhibited concerning behavior prior to the attack, including tardiness, truancy, arguments with coworkers, and poor job performance.

Insiders who committed IT sabotage held technical positions. Eighty-six percent of the insiders held technical positions. Ninety percent of them were granted system administrator or privileged system access when hired by the organization.

The majority of the insiders attacked following termination. Fifty-nine percent of the insiders were former employees, 57% did not have authorized system access at the time of the attack, and 64% used remote access. Many used privileged system access to take technical steps to set up the attack before termination. For example, insiders created a backdoor account,[3] installed and ran a password cracker,[4] took

[3] A *backdoor account* is an unauthorized account created for gaining access to a system or network known only to the person who created it.

advantage of ineffective security controls in termination processes, or exploited gaps in their organizations' access controls.

3.2 TARGETED LESSONS FOR TRAINING

Based on the above findings, the MERIT team determined that the most important lessons to be conveyed in the interactive learning environment (ILE) are the following:

Disabling access following termination is important; in order to do so effectively organizations must have full awareness of all access paths available to each of their employees. (See Section 6.4 for an explanation of access paths). Since so many acts of insider sabotage were committed following termination, the MERIT ILE must emphasize the importance of completely disabling access upon termination, a task that is often easier said than done. Many of the attacks in the Insider Threat Study were possible because the employers did not know all of the access paths available to their employees.

For example, system administrators created backdoor accounts with system administrator privileges, knowing that because account audits were not conducted the account would not be detected and would facilitate the attack following termination. Other privileged users planted logic bombs—malicious code implanted on a target system and configured to execute at a designated time or on occurrence of a specified system action. Often the insider configured the logic bomb to execute following termination, knowing that no characterization and configuration management procedures[5] were in place to detect the malicious code. Other technical insiders were able to use passwords for shared accounts because there was no formal tracking mechanism for access to those accounts. Therefore they were overlooked upon termination.

The ILE must emphasize the importance of proactive, ongoing, rigorous access management practices to facilitate complete disabling of access upon termination.

Management should carefully consider concerning behavior by an employee who appears to be disgruntled following a negative work-related event, possibly increasing monitoring of the employee's online activity. It is not practical for organizations to monitor all online activity for all employees all of the time. Determining the appropriate balance between proactive system monitoring and other duties of the IT or technical security staff is a difficult task in any organization. However, almost all insiders in the Insider Threat Study sabotage cases exhibited concerning social behavior prior to the attack. Therefore, an important lesson to be conveyed

---

[4] A *password cracker* is a program used to identify passwords to a computer or network resource; used to obtain passwords for other employee accounts.

[5] *Characterization and configuration management* refers to procedures and software that track releases and changes to software or system components so that unauthorized access can be prevented or appropriate users alerted when a file has been modified or released.

by the MERIT ILE is that organizations should maintain awareness of employee dissatisfaction and evaluate concerning behavior. Targeted monitoring of online activity by employees of concern can prevent insider sabotage through timely detection of technical precursor activity.

4 Fictional Case for Training: iAssemble, Inc.

As mentioned earlier, an interactive learning environment for training on insider threat is more effective when combined with a concrete case example that clearly illustrates the relationship between aspects of the insider threat and the effectiveness of various measures to counter the threat. However, the sensitivity of actual Insider Threat Study case data precludes the use of actual cases for training. We therefore developed a fictional case scenario that is representative of a preponderance of actual cases of insider sabotage from the Insider Threat Study.

The following characteristics of our fictional case are important:

- effective access management practices that degrade over time due to competing priorities
- increased acting out (concerning behavior) by insider
- increased tension between insider, staff, and managers
- ineffective management response that would address concerning behavior exhibited by disgruntled employee
- increased data gathering by insider
- undetected escalation of access by insider
- underestimation of insider access by management
- punitive actions that seem ineffective to management, provocative to insider

The fictional organization is called iAssemble, Inc.[6] The full text of the iAssemble case example appears in Appendix A. A summary of the case follows:

iAssemble sold computer systems directly to customers; building each system made-to-order at competitive prices. Ian Archer, the insider threat actor, had been with iAssemble since its founding and was the sole system administrator. The environment at iAssemble was traditionally very relaxed. However, recent substantial company growth resulted in a change in culture, as well as new management who hired a new lead system administrator over Archer.

This action triggered Archer's disgruntlement; he felt his hard work over the years was not appreciated. In addition, the new lead system administrator restricted the privileges of all iAssemble employees, including Archer. Archer vented his anger by openly harassing individuals and purposely stalling progress on important projects. A performance improvement plan was instituted by Archer's new manager with disciplinary actions including written warnings, a temporary suspension, and reduction in his salary. Suspecting he would soon be fired, Archer created a back-

[6] The iAssemble organization and case example are completely fictional; any resemblance to a real organization or insider threat case is unintentional.

door with system administrator privileges on iAssemble's server for later access should his authorized access be disabled or his administrative privileges be revoked.

Management's increased sense of risk of malicious activity by Archer led them to ramp up audits of access control quality and access management. Unfortunately these measures were put in place too late to prevent or detect Archer's backdoor installation. When management fired Archer they disabled all known access paths. But unknown to management, a coworker had shared his password with Archer to increase productivity for their project team. Archer used that password to log in remotely to the coworker's machine the night of his firing. Using the backdoor account he installed a logic bomb on the machinery server, set to detonate three months later.

The logic bomb deleted all the files on the machinery and backup servers leaving the assembly lines at iAssemble frozen. An investigation revealed that access control policies and practices had eroded over time. The investigation lead to the arrest of Ian Archer, but iAssemble was left on shaky ground, causing share prices to plummet. Their image in the market was blemished and stockholders demanded detailed explanations from company management.

iAssemble's decision to increase monitoring and auditing was the right one—the steps they took increased management's knowledge of employees' access paths. However, the gap between management's knowledge of Archer's access paths and his actual access had not been fully eliminated when he was fired, so iAssemble could not disable all of his access paths in time. Hence, Archer was able take advantage of the residual access that he had to attack following termination.

We believe that the iAssemble case provides a coherent and well-grounded basis for training on the access management issues relevant to insider sabotage and is representative in character (but not necessarily detail) of many of the actual cases that we have seen.

5 Methodological Background

As mentioned previously, our effort towards modeling insider sabotage activity uses a technique called system dynamics (Sterman 00). System dynamics is a method for modeling and analyzing the holistic behavior of complex problems as they evolve over time. System dynamics has been used to gain insight into some of the most challenging strategy questions facing businesses and government for several decades. The Franz Edelman Prize for excellence in management was given in 2001 to a team at General Motors who used system dynamics to develop a successful strategy for launch of the OnStar System (Huber 02). System dynamics is particularly useful for gaining insight into difficult management situations in which our best efforts to solve a problem actually make it worse. Real problematic situations in which system dynamics helps create clarity include the following (Sterman 00):

- Efforts to build new roads to alleviate traffic congestion only result in increased congestion.
- Use of cheaper drugs pushes costs up, not down.
- Lowering the nicotine in cigarettes, supposedly to the benefit of smoker's health, only results in people's smoking more cigarettes and taking longer, deeper drags to meet their nicotine needs.
- Levee and dam construction to control floods leads to more severe flooding by preventing the natural dissipation of excess water in flood plains.
- Applying more resources to incident response to handle a high workload takes resources from proactive management activities and increases the incident workload.

Our application of system dynamics targets insider sabotage of an organization's information, systems, or networks. Intuitive solutions to problems in this area often reduce the problem in the short term, but make it much worse in the long term. System dynamics is a valuable analysis tool for gaining insight into solutions that are effective over the long term and for demonstrating their benefits.

A powerful tenet of system dynamics is that the dynamic complexity of problematic behavior is captured by the underlying feedback structure of that behavior. So we decompose the causal structure of the problematic behavior into its feedback loops to understand which loop is strongest (i.e., which loop's influence on behavior dominates all others) at particular points through time. We can then thoroughly understand and communicate the nature of the problematic behavior and the benefits of alternative mitigations.

System dynamics model boundaries are drawn so that all the enterprise elements necessary to generate and understand problematic behavior are contained within them. This approach encourages the inclusion of soft (as well as hard) factors in the

model, such as policy, procedural, administrative, or cultural factors. The exclusion of soft factors essentially treats their influence as negligible, which is often not the case. The inclusive, endogenous viewpoint helps show the benefits of mitigations to the problematic behavior that are often overlooked by low performers, partly due to their narrow focus on technical solutions to resolve problems.

We rely on system dynamics as a tool to help test the effect of strategies for improving the performance of IT management. In some sense the simulation of the model will help predict the effect of these strategies. But what is the nature of the types of predictions that system dynamics facilitates? Dennis Meadows offers a concise answer by categorizing outputs from models as described below (Meadows 74):

1. absolute and precise predictions. (Exactly when and where will the next cyber attack take place?)
2. conditional precise predictions. (If a cyber attack occurs, how much will it cost my organization?)
3. conditional imprecise projections of dynamic behavior modes. (If a bank mandates background checks for all new employees, will its damages from insider fraud be less than they would have been otherwise?
4. current trends that may influence future behavior. (If the current trends in insider IT sabotage continue, what effect will this have on my company's performance in five years?)
5. philosophical explorations of the consequences of a set of assumptions, without regard for the real-world accuracy or usefulness of those assumptions. (If a telepathic alien race invaded earth, how would this affect my risk of insider attack?

The model we develop, and system dynamics models in general, provide information of the third sort. Meadows explains further that "this level of knowledge is less satisfactory than a perfect, precise prediction would be, but it is still a significant advance over the level of understanding permitted by current mental models."

5.1 NOTATION

In graphic representations of the model we describe, signed arrows represent the system interactions, where the sign indicates the pair-wise influence of the variable at the source of the arrow on the variable at the target of the arrow:

- A positive (+) influence indicates that if the value of the source variable increases, then the value of the target variable increases above what it otherwise would have been, all other things being equal. And, if the value of the source variable decreases, then the value of the target variable decreases below what it would otherwise have been, all other things being equal.
- A negative (-) influence indicates that if the value of the source variable increases, then the value of the target variable decreases below what it would otherwise have been, all other things being equal. And, if the value of the

source variable decreases, then the value of the target variable increases above what it would otherwise have been, all other things being equal.

We can illustrate the above definitions using the influence diagram shown in Figure 1, which represents a very simple room heating system. A positive influence is indicated by the arrow from *rate of heat input* to *room temperature*. At a particular thermostat setting, as the rate of heat input increases (or decreases), then the temperature of the room increases (or decreases) above (or below) what it would have been. A negative influence is indicated by the arrow in the other direction. As the room temperature increases (or decreases), the rate of heat input decreases (or increases) below (or above) what it would have been, as would be expected by a room heating system.

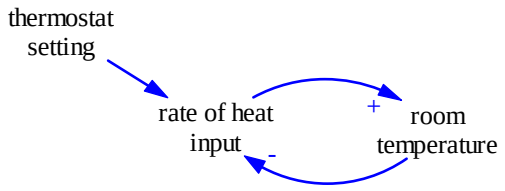


Figure 1: A Simple Feedback Loop

As mentioned previously, dynamically complex problems can often be best understood in terms of the feedback loops underlying those problems. There are two types of feedback loops: balancing and reinforcing. Balancing loops describe aspects of the system that oppose change, seeking to drive organizational variables to some goal state. Reinforcing loops describe system aspects that tend to drive variable values consistently upward or consistently downward. The polarity of a feedback loop is determined by "multiplying" the signs along the path of the loop. Balancing loops have negative polarity and reinforcing loops have positive polarity.

Figure 1 depicts a balancing loop that seeks to move the room temperature to the thermostat setting. This system is balancing as shown by the odd number of negative signs along its path. The goal state is a room temperature equal to the thermostat setting. In general, balancing loops describe aspects that oppose change, and usually involve self-regulation through adaptation to external influences.

Figure 2 shows a more interesting example in the domain of project management. Figure 2a depicts one approach an organization may adopt in trying to put a project that is behind schedule back on track: having its employees work overtime. The closed form in Figure 2b shows the corresponding balancing feedback loop that characterizes the goal of the approach as moving the project to the state of being on schedule.

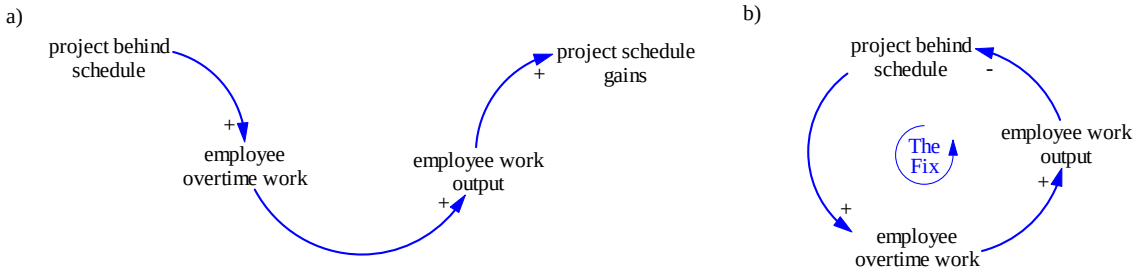


Figure 2: (a) Project Management—Desire to Use Overtime to Correct Schedule; (b) Closed-Loop Representation Showing (Balancing) Feedback to Improve Progress

Figure 3 shows that the project-management behavior described above is subject to a reinforcing feedback loop in which overtime in the long term leads to employee burn out, lower quality of work, and the need to rework defective artifacts. The longer this goes on the further the project gets behind schedule because of the increasing amount of rework. This combines with the previous balancing feedback loop, where the balancing loop dominates in the near term with the reinforcing loop taking over with increasing amounts of employee overtime and burnout. This type of thinking about the feedback structure of systems and about which feedback loops dominate at different periods in time is characteristic of system dynamics modeling and analysis.

The reinforcing loop is shown mostly in red but it shares part of the influence path of the balancing loop from *project behind schedule* to *employee overtime work*. The reinforcing nature of the feedback loop is evident from the even number of negative signs along its path.[7] Reinforcing loops may help explain explosive growth or implosive collapse of a system.

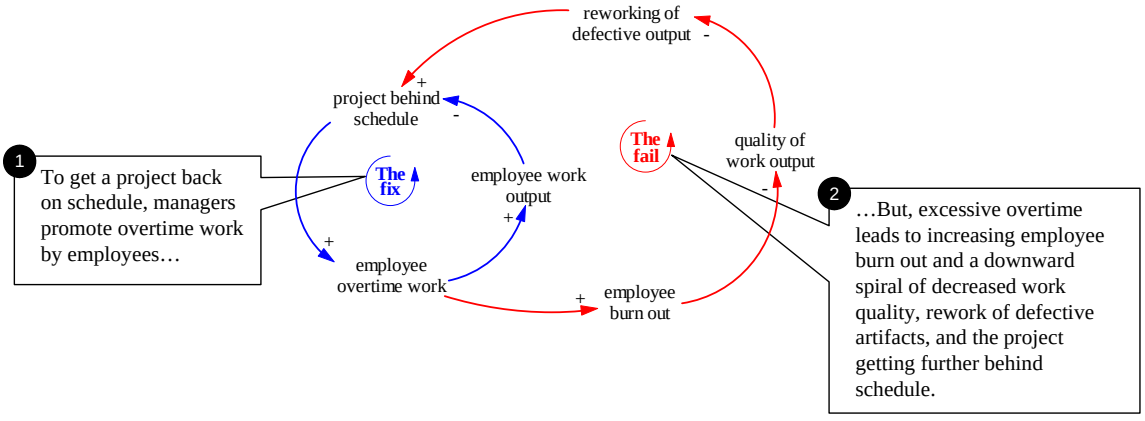


Figure 3: Unintended Burnout Due to Overtime

[7] Feedback loops that have no negative signs along the influence path have positive polarity and thus are reinforcing loops.

A quantitative system dynamics model refines and describes the relationships in the qualitative system dynamics model using mathematical equations. This is done by adding two new concepts to the modeling notation: stocks and flows.

- Stocks represent accumulations of physical or non-physical quantities and flows represent the movement of these quantities between stocks. Stocks are depicted as named boxes within the model.
- Flows are depicted as double-lined arrows between the stocks with a named valve symbol indicating the name of the flow. Flows that come from (or go to) a cloud symbol indicate that the stock from which the flow originates (or to which the flow goes) is outside the scope of the model.

The next section describes the assumptions on which our model is based. The remainder of the paper describes the stock and flow model and the simulation results from that model.

6 Model Assumptions

The following assumptions are key to understanding the dynamics, relationships, and conditions in the MERIT model related to insider IT sabotage and its countermeasures.

6.1 SCOPE

The model begins with the insider at his or her highest position in the organization. Most insiders did not carry over problems from positions prior to their highest positions. While the time periods related to firing and demotion are important in the model, hiring is not, because significant triggering events related to the attack or prompting the desire to attack did not typically occur soon after being hired. Also within the model scope is a negative work-related event that causes the insider to feel dissatisfied toward the organization, supervisors, or co-workers. Termination or demotion is frequently the last negative work-related event triggering the insider attack.

6.2 ORGANIZATION

Key to the model is the organization's knowledge of insider access rights and privileges, rather than the authorization for and legitimacy of insider actions. Most organizations have the ability to track, monitor, and identify access paths for employees, but they can either be unaware of or forget access paths available to employees due to poor security management practices. In addition, practices such as security awareness and education, account management, and personnel behavioral management, play an important role in the model. Access control is key because insiders require access to perform their job functions but can also use this access to attack. Therefore, imperfect states of practice, particularly with regard to access control, have a heavy influence in the model. Access control may not be perfect at the start of the simulation.

6.3 INSIDERS

The next group of assumptions deals with the insiders themselves: their means, motives, and opportunities. The model assumes that insiders work alone in attacking the organization and do not collude; this assumption is supported by most of the cases examined in the Insider Threat Study. The method by which the insider attempts to attack the organization is typically limited to the skill sets, experiences, and education exhibited while still an employee with the organization. Insiders may attempt or succeed at gaining more access than their organizations authorize, but they will seek to gain this access within the confines of their current skill sets, experience, and education.

Insiders tend to feel entitled to perform certain actions or act in a specific way, and this entitlement escalates over time. If they do not receive reprimands, sanctions, or correction, insiders begin to feel that they have the organization's authorization to behave irregularly. When insiders are penalized or corrected they may react negatively and cause further behavioral disruption or commit technical sabotage.

6.4 ACCESS

Ironically, while access is granted as a necessary course of conducting business operations, it is also one of the most essential elements of insider attacks. Access to information and systems allows employees to read, modify, and delete business and system data. The following section expands on employee access paths that are frequently used to conduct attacks.

6.4.1 Access paths

In the MERIT model, access is provided through *access paths*: a set of one or more access points leading to a critical system. Examples of points along access paths are employee badges, computer accounts, passwords, and Virtual Private Networks (VPN). The model presumes that the insider obtains access in one of three ways—access paths are granted by the organization, created by the insider, or discovered by the insider. Access paths can be known or unknown to the organization. An access path that is unknown to management is not necessarily illegitimate, but organizations should reduce unknown access paths by identifying them, reviewing each for validity, and disabling those that fulfill no justified business need.

Granted paths are those authorized by the organization. For example, a granted access path for a web server administrator could be software and hardware used to publish corporate web pages. One problem with granted access paths, illustrated in the model, is that organizations can lose track of their existence if formal tracking procedures are not enforced. An example of a forgotten path is a privileged shared account created for a team of software developers for the duration of a project that is not removed or restricted after the project terminates.

Created paths are those established by an employee, such as computer accounts that are created or hacking tools that are installed on the system by the insider. Created paths can be authorized or unauthorized, and the organization may or may not know of their existence.

Discovered paths are existing paths revealed to or discovered by an employee. Although they can be used for malicious insider actions, they may not have been created with malicious intent. Discovered access paths are those found when employees learn that they can access information, resources, or network services they did not know existed or for which they did not know they had legitimate access.

Other access path assumptions in the model include the following:

- Insiders can lose some or all of their access paths, as well as the ability to create new paths.
- Insiders who are demoted or terminated may retain the ability to create or use access paths for which they are no longer authorized because of a lapse in procedure or practice.
- It takes time for organizations to recover from poor access management.
- Effective disabling of access paths requires that management have full awareness of all paths available to the employee and of the employee's ability to create new paths.
- Even without deliberate action by an insider to obtain a higher level of access, there tends to be a gradual increase in the number of access paths available to an insider over time.

6.5 DEFENSES

The final assumption pertaining to the model deals with organizational defenses and responses to unacceptable employee behavior. Organizations typically use administrative, physical, and technical controls to deter, prevent, detect, and respond to attacks on information and systems, including insider attacks. The MERIT model focuses on administrative and technical controls, since physical controls were not a predominant factor in most cases in the Insider Threat Study.

Administrative and technical controls relevant to mitigating risk of insider threat in the MERIT model are described in Table 1.

Table 1: Simulation Effects of Policy Levers

Policy Lever	Description	Effect
employee intervention	Positive interventions like employee assistance or counseling that attempt to lower disgruntlement directly, to reduce inappropriate behavioral or technical actions by insider.	May not be effective if quality of intervention is low.
sanctioning	Punitive measures that attempt to motivate the insider to reduce his inappropriate behavioral or technical actions to avoid additional sanctioning.	May have the opposite effect of increasing disgruntlement and inappropriate actions.
technical monitoring	Real-time measures to track and analyze an insider's online actions, such as the use of access paths or information and resources accessed.	If technical monitoring is not initiated or quality is low, management may not have an accurate sense of the risk that an insider poses to the organization.

training	Currently limited to education of employees on appropriate usage of computer and network systems and the consequences if misused.	Training quality affects the rate of inappropriate online actions and attacks by insiders.
tracking	Efforts by management to keep track of access paths.	Poor tracking leads to high rates of access paths unknown by management, making it more difficult to disable paths and easier for the insider to conceal his actions.
auditing and disabling access paths	Efforts by management to discover, understand, review, and disable access paths available to the insider. Enables comparison between acceptable policies and procedures and employees' abilities and efforts to access information, create access paths, or use access paths.	Facilitates discovery of access paths available to the insider. Poor audit allows insiders to amass many unknown access paths making it easier to conceal actions and attack after termination.
termination threshold	The threshold of risk posed by the insider to the organization above which management fires the insider.	Too high a threshold may give a malicious insider additional time to attack the organization or take technical actions to set up an attack following termination. Too low a threshold may cause the organization to terminate valuable employees who just need a little intervention to solve their problems.
termination time	The time it takes the organization to terminate an insider once the *termination threshold* is reached.	If termination time is too long then the insider may maintain authorized access to the system long enough to facilitate an attack.
employee intervention	Positive interventions like employee assistance or counseling that attempt to lower disgruntlement directly, to reduce inappropriate behavioral or technical actions by insider.	May not be effective if quality of intervention is low.

7 Modeling Behavioral Aspects

Employee disgruntlement was a recurring factor in the Insider Threat Study sabotage cases, predominately due to some unmet expectation by the insider. This is evident in the three examples below:

1. The insider expected certain technical freedoms in his[8] use of the organization's computer and network systems, such as storing personal MP3 files, but was reprimanded by management for exercising those freedoms.
2. The insider expected to have control over the organization's computer and network system, but that control was revoked or never initially granted.
3. The insider expected recognition or prestige from management, but was disturbed upon some event in the workplace, such as being passed over for a promotion.

In our model we focus on the first two. Insider freedom thus represents freedom for the insider to use or control the system. Expected freedoms could be measured either by the number or extent of privileges or on a continuous scale from none to root access.

7.1 INSIDER EXPECTATION OF FREEDOM

Figure 4 depicts changes in the insider's expectations over time based on his actual freedom as well as the insider's *predisposition to disgruntlement*. This predisposition differs from one person to the next, and influences the rate that expectations rise and fall. The rise of expectations is influenced heavily by the *actual freedom given insider*. As illustrated in reinforcing loop R1, with lax management controls actual freedom grows commensurately with expected freedoms. As more freedom is allowed, more freedom is taken; as more freedom is taken, more is allowed. In the model, it is assumed that even lax management sets an upper bound on the extent of freedoms allowed to any employee.

Lax management unintentionally encourages escalation of expectation. Expectation escalation is seen in the simulation results in Figure 5. The simulation starts off with expected and actual freedom at an equal value of 10 on a scale of *relative freedom.* This is a rather arbitrary measure of the relative freedom allowed any employee of the organization according to the organization's appropriate systems usage policy. With lax management, some employees will try to "push the envelope," using the system as desired regardless of the organization's usage policy. This is especially true for insiders with a strong sense of entitlement.

[8] Ninety-six percent of the insiders in the Insider Threat Study who committed IT sabotage were male. Therefore, male gender is used to describe the generic insider throughout this technical note.

As management allows the insider's actual freedom to increase beyond that permitted by policy, the insider's expectation also rises. As shown in the figure, expected and actual freedom continue to increase at an equal rate until about week 40, when freedom reaches a point that even lax management will not permit—more than twice the freedom allowed by policy. At this point, the insider expects slightly more than is permitted; this situation creates an equilibrium condition where *unmet expectation* stays fairly constant over time.

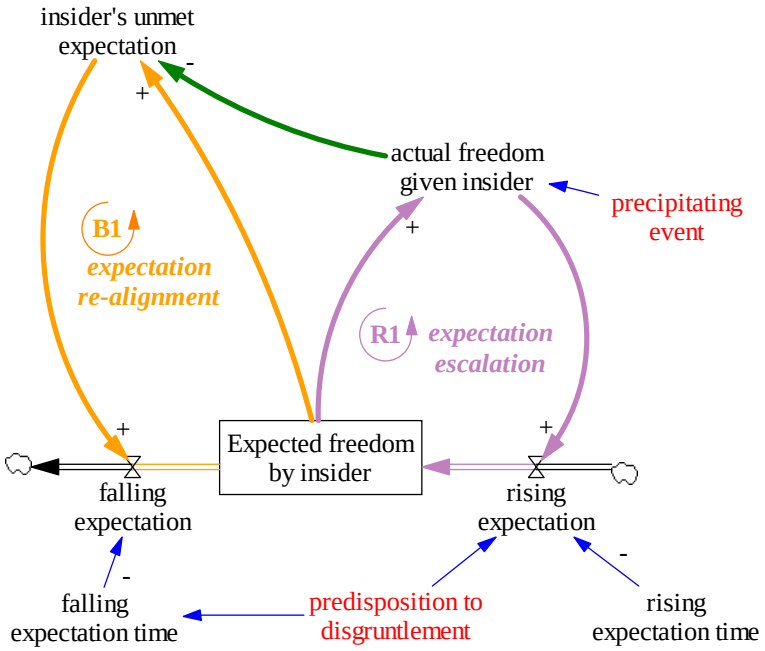


Figure 4: Expected Freedom by Insider

This simulation illustrates a situation in which lax management permits increasing freedom for the insider that can cause major problems later on, especially if that insider has a predisposition for disgruntlement. The trigger for those major problems, which we call the *precipitating event*, tends to be anything that removes or restricts the freedom to which the insider has become accustomed. In the iAssemble case, as in some of the cases in the Insider Threat Study, the trigger is the hiring of a new supervisor who enforces the organization's system usage policy.

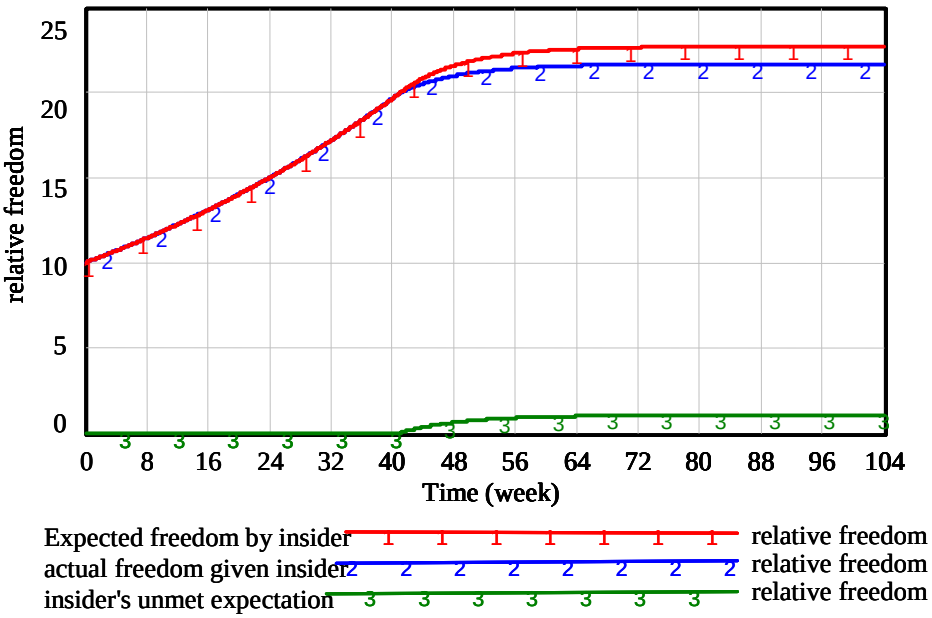


Figure 5: Expected and Actual Freedom Growth with Lax Supervisor

Figure 6 shows simulation results with a new supervisor hired at week 20 who enforces the usage policy, shown by the drop in *actual freedom given insider* to 10, the *relative freedom* of an insider abiding by that policy. Coincident with the drop is a commensurate rise in *unmet expectation*. Expectation rises (about 40% in 20 weeks) much faster than it falls, approaching its original policy level at around week 92, assuming an insider with a strong sense of entitlement. Barring any additional loss of freedom, however, expectations do fall gradually as the insider comes to accept his new situation. Nevertheless, the period of high unmet expectation is one of high risk for the organization as explained below. The additional drop in the *actual freedom given insider* is also explained below.

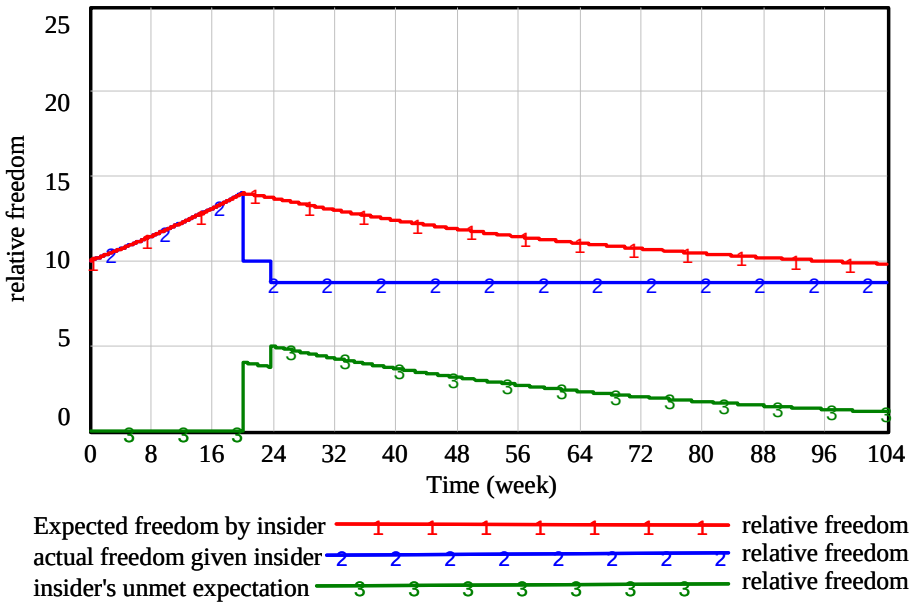


Figure 6: Expected and Actual Freedom with Strict Supervisor Hired at Week 20

7.2 ESCALATION OF DISGRUNTLEMENT AND SANCTIONING

Figure 7 depicts part of the model, influences of *unmet expectation* on the insider's offline[9] behavior, and the organization's response. Three additional stocks are introduced:

1. *Insider disgruntlement*: the insider's internal feelings of discontent due to demands or restrictions by the organization that he perceives as unacceptable or unfair.
2. *Behavioral precursors*: observable aspects of the insider's offline/social behavior inside or outside the workplace that might be deemed inappropriate or disruptive in some way.
3. *Sanctions*: the organization's punitive response to inappropriate behaviors. Sanctions can be technical, such as restricting system privileges or right to use the organization's equipment at home, or non-technical, such as demotion or formal reprimand.

A generic measure of *relative severity* is used to measure behavioral precursors, damage, and disgruntlement.

Reinforcing loop R2 in Figure 4 characterizes escalation of disgruntlement in response to sanctions for inappropriate social behaviors. As the *insider's unmet expectations* increase, *Insider disgruntlement* increases. Insiders exhibit disgruntlement by *acting inappropriately offline*. Observable inappropriate offline behaviors

[9] Throughout this technical note, *online behavior* refers to actions taken using the computer, while *offline behavior* refers to social behaviors that are not taken on the computer.

vary; some insiders take revenge primarily online, exhibiting fewer offline precursors. We assume that the insider's *predisposition to disgruntlement* indicates his tendency to engage in inappropriate offline behavior before an attack.

Continuing around loop R2 of Figure 7, notice that *Severity of the actions perceived by org* is affected by *time to realize insider responsible*.[10] Severity of actions influences the extent of sanctioning, which further limits the *actual freedom given insider*. These dynamics explain the second decrease in *actual freedom* in Figure 6 at around week 24, after the new supervisor imposes sanctions further limiting the insider's freedom. The model in Appendix B also shows that technical restrictions on the insider can further limit the insider's actual freedom.

Instead of (or in addition to) punitive measures, organizations may take positive actions to address an insider's disgruntlement. Such actions, represented as *employee intervention*, include referral to an employee assistance program or counseling. Balancing loop B2 in Figure 7 reflects use of *employee intervention* to address disgruntlement. The organization's perception of the severity of the *Behavioral precursors*, the observable manifestation of the insider's disgruntlement, and organizational policies determine whether positive intervention or sanctions are warranted.

[10] *Severity of actions perceived by org* is a smooth of several factors and may also be considered a stock.

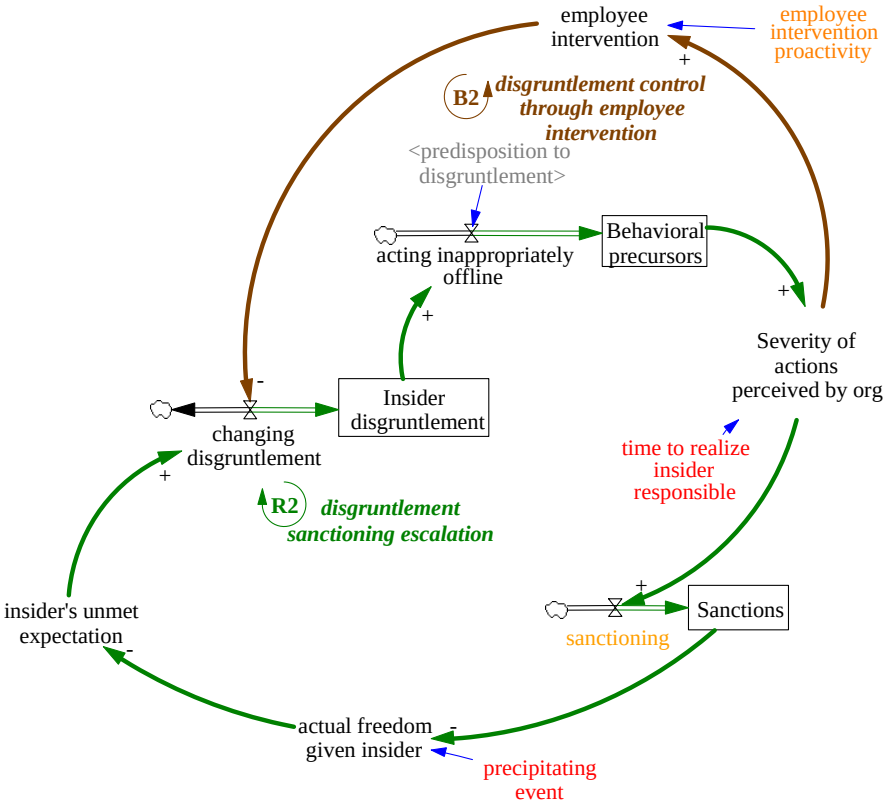


Figure 7: Escalation of Disgruntlement and Sanctioning

Figure 8 shows the increase of *Insider disgruntlement* due to the *insider's unmet expectation* that arises due to the new supervisor's strict enforcement of the organization's usage policy. With only minimal *employee intervention* (0.2 on a scale from 0 to 1), disgruntlement rises to almost three times its normal level at about week 24. The predisposed insider begins to act out offline and receives two sanctions during this time period.

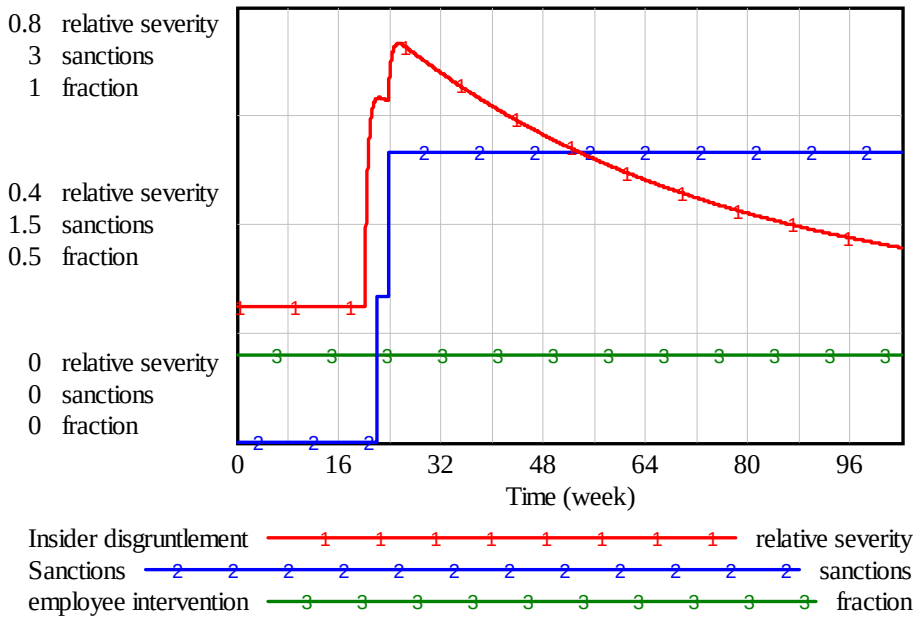


Figure 8: Escalation of Disgruntlement and Sanctioning with Minimal Intervention

Figure 9 provides a notional view of how proactive employee intervention can decrease both disgruntlement and the sanctions needed to address inappropriate behavior arising from that disgruntlement. In this case, even the predisposed insider is much less disgruntled and warrants less of a punitive response, i.e. only one sanction. One nice aspect of employee intervention is that by treating disgruntlement directly, there is less need for punishment and corresponding less disgruntlement caused by the punishment. Thus, when intervention works it is a win-win situation for both the organization and its employees. We are still investigating the general characteristics of the insider and the intervention itself that underlie the success of the approach.

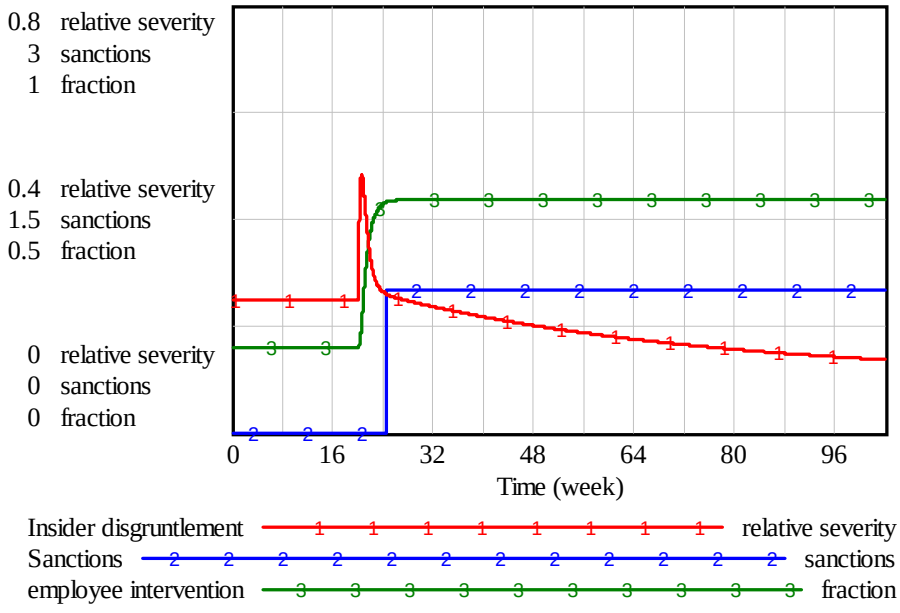


Figure 9: Disgruntlement and Sanctioning with Proactive Intervention

8 Modeling Technical Attack Aspects

As previously mentioned, an organization's full awareness of access paths available to an insider is critical to being able to disable those access paths when needed. Two stocks model this dependency: *Insider access paths unknown to org* and *Insider access paths known to org*.

Figure 10 shows the flows between these two stocks:

- *forgetting paths* flow : Management or the IT staff may forget about known paths, making them unknown. For example, a manager might authorize a software developer's request for the system administrator password during a time of heavy development, but if a formal list of employees with access to that password is not maintained then the manager could forget that decision over time, or the manager could leave the organization, leaving no "organizational memory" of the decision.
- *discovering paths* flow: Management or the IT staff can discover unknown paths, making them known. Discovery can be accomplished by auditing, for example, when new accounts could be discovered with system administrator or privileged access that were previously unknown to management.

Insiders can acquire new paths unknown to the organization via the *acquiring unknown paths* flow. Finally, organizations can disable known paths via the *disabling known paths* flow. This stock and flow structure is used in the refinement of the technical aspects of the model described in the section below.

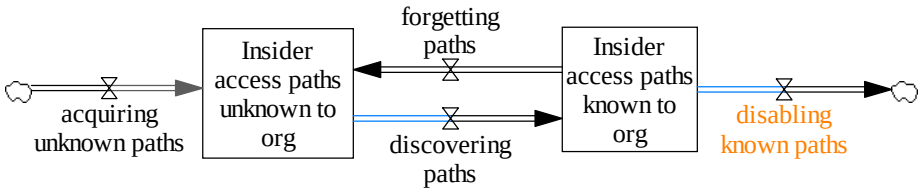


Figure 10: Access Path Stocks and Flows

8.1 ATTACK SETUP AND CONCEALMENT

As discussed earlier, an insider's *predisposition to disgruntlement* and unmet expectations can lead to increasing disgruntlement that, if left unchecked, can spur not only behavioral precursors but technical disruptions and attacks on the organization's computer and network systems. Prior to the actual attack, there are typically *Technical precursors*—actions by the insider to either set up the attack (for example, installation of a logic bomb) or to put in place mechanisms to facilitate a future attack (for example, creation of backdoor accounts to be used later for the attack). Such an online *Technical precursor* could serve as an indicator of a pending attack

if detected by the organization. Figure 11 depicts the influence that insider disgruntlement can have on the occurrence of *Technical precursors* that could indicate a pending attack. The figure shows that both unknown and known access paths can be used to set the stage for attack.

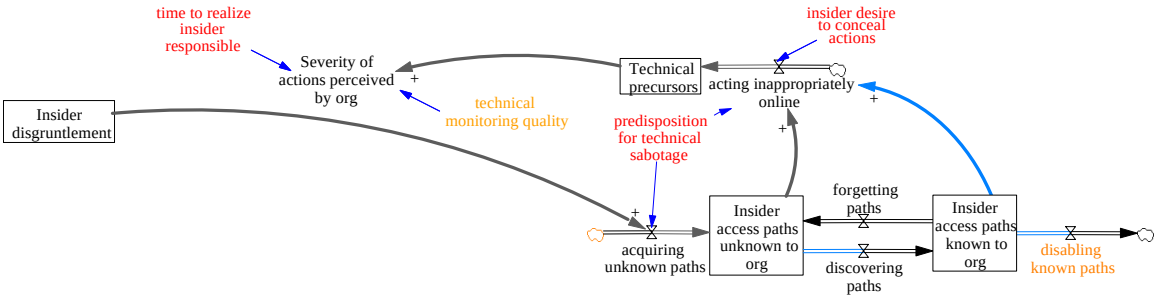


Figure 11: Attack Setup and Concealment

The extent to which insiders rely on unknown access paths depends on their *desire to conceal actions*. Insiders who do not care whether they are caught, or insiders acting impulsively (often out of the passion of the moment), may use both known and unknown paths in their attack. Insiders who are particularly risk averse may only attack using access paths that are unknown to the organization. Of course, an insider may not know whether the organization is aware of a particular access path or not. Nevertheless, in either case, insiders generate *Technical precursors* that suggest suspicious activity. To perceive the severity of these precursors, the organization must have a *technical monitoring quality* sufficient to detect the precursors in the first place.

8.2 ATTACK ESCALATION

As shown in Figure 12, *Insider disgruntlement* contributes directly to the rate of inappropriate technical actions taken by the insider, especially actions that facilitate the attack. Some of these actions also contribute to the damage potential of the attack. Examples include sabotage of backups and decreases in the redundancy of critical services or software.

Since insiders in most sabotage cases studied were motivated by revenge, the model assumes that the actual attack occurs once the damage potential reaches an *attack threshold* defined by the insider, provided that the disgruntlement level is sufficiently high. Multiple attacks may be executed provided that a sufficient number of access paths are available to set up and execute the subsequent attacks. If the attack execution is autonomous (for example a logic bomb set to go off when the system reaches a certain state) the insider may need no access paths to the organization's critical systems in order to execute the attack. In such a case, the planting of the logic bomb could actually be considered to be the attack.

Figure 13 shows the simulation results with *predisposition for technical sabotage* and *insider desire to conceal actions* set to 1. *Insider disgruntlement* rises to its highest level at about week 35 after which the attack is executed. At week 20, prior

to the attack, *Behavioral precursors* are the first observables indicating the insider's disgruntlement. About 4 weeks later the *Technical precursors* start to appear. We believe that this pattern of behavioral disruption before technical disruption is common among many insiders.

Some insiders install unauthorized tools for non-malicious purposes immediately after they are hired. For example, one system administrator installed a root kit—a "hacker" tool used to maintain access to a system without the owner's knowledge. These tools provide insiders with unauthorized access that makes their work more convenient. These actions are technical precursors that increase risk of an attack if that insider ever becomes angry enough to take revenge on his employer.

The severity of the technical precursors rises above the severity of the behavioral precursors at about the time of the attack. *Attack damage*, which for obvious reasons has severity much higher than the behavior or technical precursors, occurs immediately at the time of the attack. Technical precursors level off immediately after the attacks because disgruntlement is greatly reduced by attack execution. We are still investigating the exact relationship between attack execution and insider disgruntlement, but we believe that the behavior exhibited by the current model to be representative of the insider sabotage cases in our study.

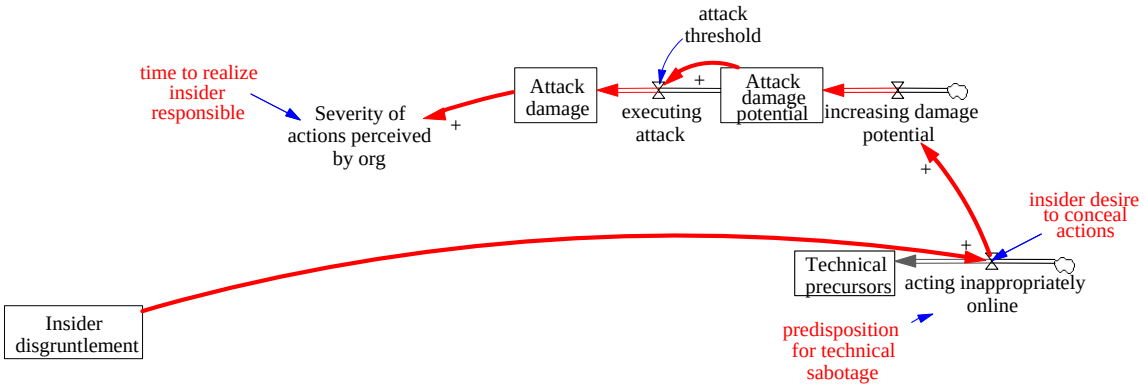


Figure 12: Attack Escalation

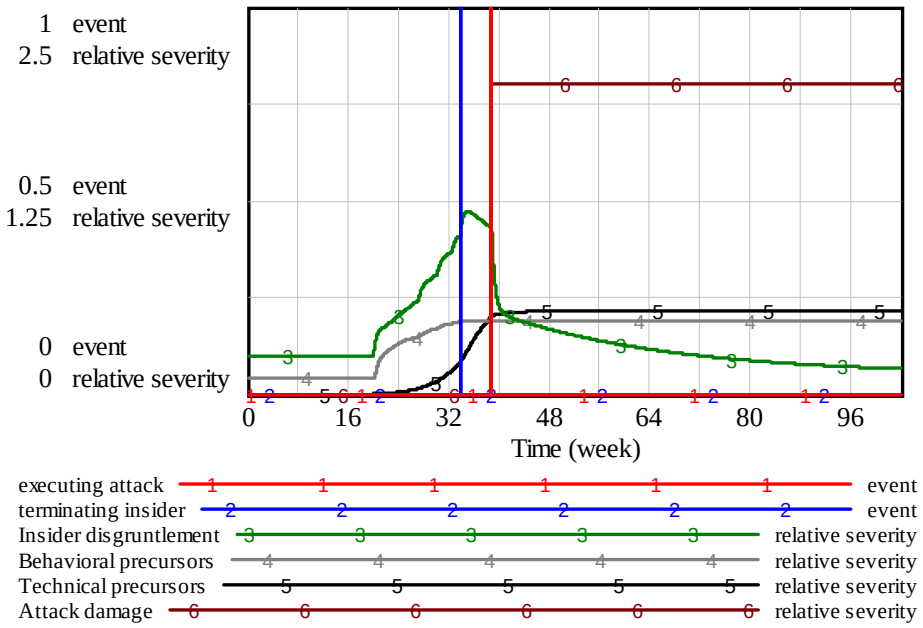


Figure 13: Attack Simulation

9 Modeling Technical Defense Aspects

In the fictional case, iAssemble's defenses against insider attack were purely reactive, based on severity of insider actions and risk subsequently perceived. Figure 14 depicts two defensive actions:

1. *Auditing the organization's systems to discover unknown access paths available to the insider*: Auditing must be followed by disabling those paths (loop B3) for this defense to have significant effect. It is possible, however, that an organization's discovery of access paths would be a sufficient deterrent for a risk-averse insider if the insider knew the organization had discovered the paths.
2. *Reducing the insider's access path creation ability (loop B4)*: This defense reduces the insider's ability to acquire new unknown paths.

Both of these defenses target access paths that may be used by the insider to set up or execute an attack. A risk-averse insider, who will not attack unless he can conceal his actions, will have less incentive to attack if unknown access paths are disabled. Known access paths can also be disabled if they are not needed by the insider to perform critical job functions.

If an organization disables the access paths required to fulfill the insider's job responsibilities, his performance and the organizational mission may be negatively affected. However, if the perceived risk is sufficiently high, the organization may choose to disable the insider's access paths anyway. Within the simulation run, if the risk reaches a termination threshold, the insider is fired and all known access paths are immediately disabled. Of course, any unknown access paths still available to the insider may be used to set up and execute an attack.

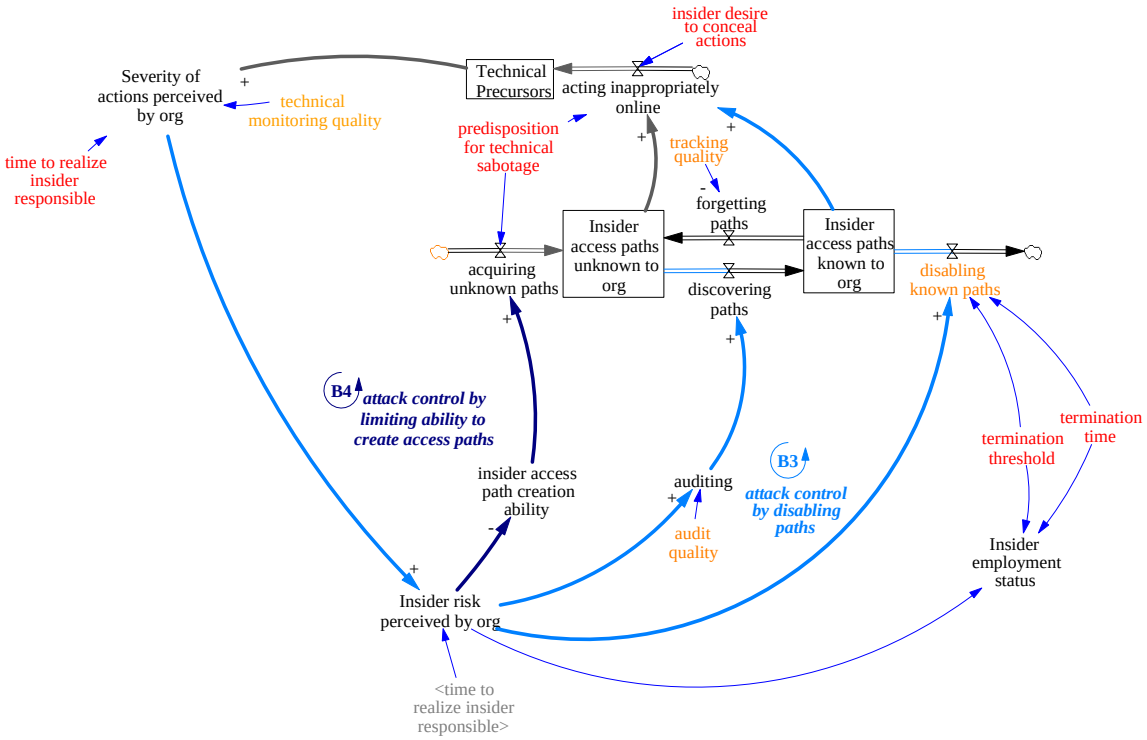


Figure 14: Risk-Based Auditing and Access Path Disabling

Figure 15 shows the results of executing the model with audit quality set at 50%, the same level of audit used to generate the results for Figure 13. In Figure 15, however, the number of access paths available to the insider, both known and unknown to the organization, is shown over time. The insider is terminated at about week 32; all known access paths are disabled at time of termination. However, with audit quality at 50%, the insider still has enough unknown access paths to continue to set up and execute the attack at about week 38.

To test the effects of auditing, Figure 16 shows the results using the same parameter settings except that *audit quality* is set to 80%. Here, the higher level of audit quality keeps the number of access paths unknown to the organization sufficiently low so that no attack can be executed, before or after termination. The *Technical precursors* suggest that the insider started to set up the attack, but the organization's defenses were sufficient to stop the insider before he reached the *attack threshold.* For an attack threshold of severity 2, the tipping point for the attack is an audit quality of between 68% and 69%. Future work will determine what it means for an audit process to be of a certain quality. Of particular interest will be the characteristics of an audit at the tipping point for an attack.

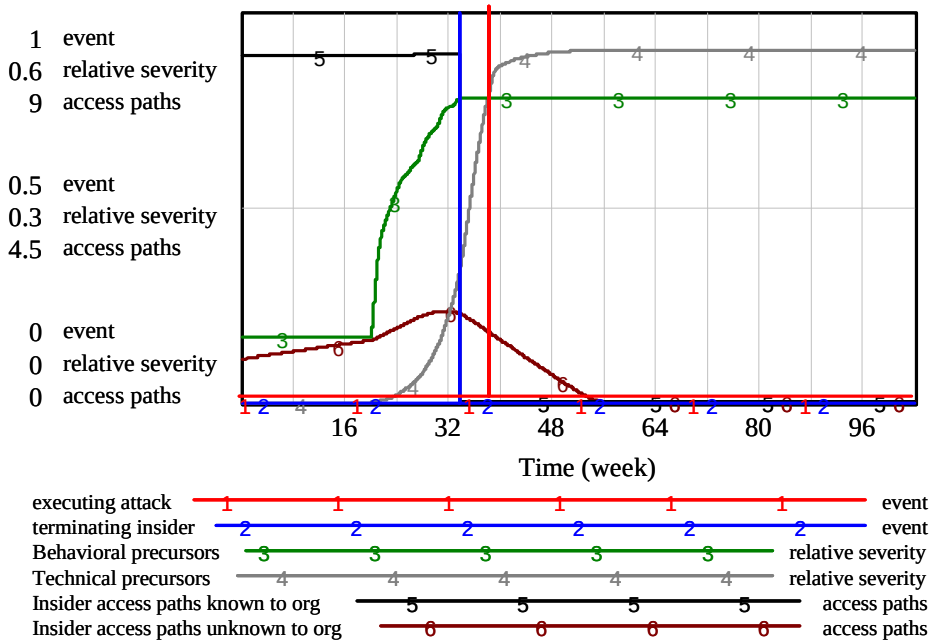


Figure 15: Attack Simulation with Audit Quality at 50%

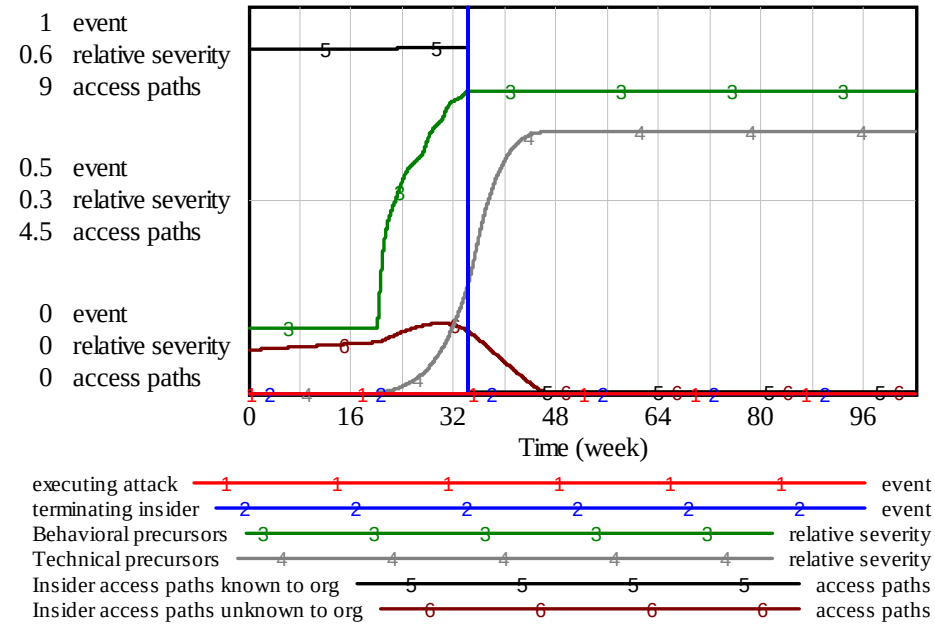


Figure 16: Attack Simulation with Audit Quality at 80%

10 Exhibiting the iAssemble Reference Mode

This section demonstrates that the MERIT model exhibits the behavior of the iAssemble case. Since the iAssemble case is representative of a preponderance of sabotage cases analyzed in the Insider Threat Study, we believe this model represents key issues in insider IT sabotage cases. From this we infer that the model is useful for identifying and analyzing the solution space, which includes policy, procedural, and technical measures that, when used together, can significantly help prevent or detect insider IT sabotage.

Figure 17 shows the increasing gap between the perception of the insider's access paths and the actual access paths available to him, which has the same general pattern as in the iAssemble case of Figure 22.

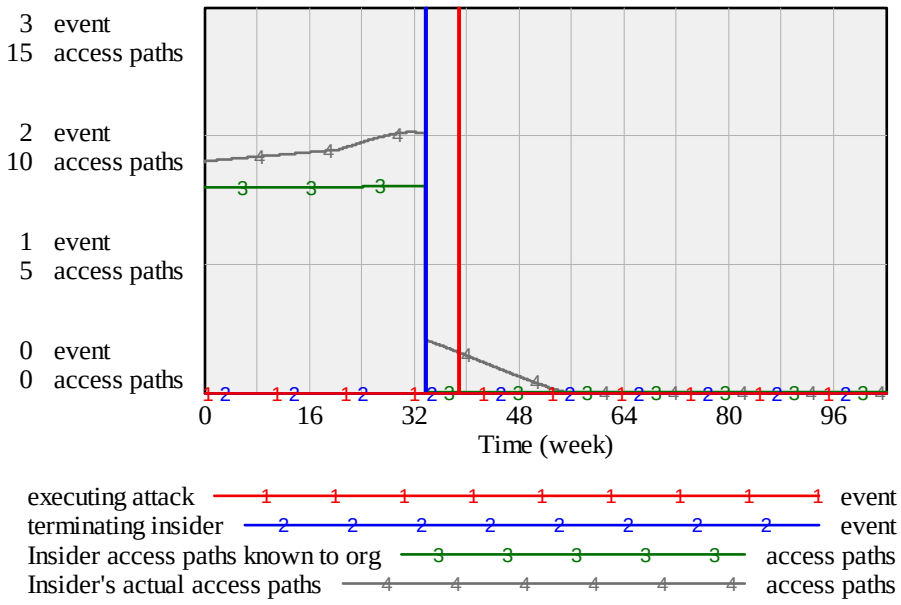


Figure 17: Exhibiting the iAssemble Problematic Behavior

This perception gap indicates an erosion of the organization's control of access to its systems. *Access control quality* (ACQ) is defined as follows:

access control quality

= W_u x ***unknown path access control quality***

\+ W_k x ***known path access control quality***

where

- w_u is the weight that the organization gives to unknown access paths to determine the access control quality

- w_k is the weight that the organization gives to known access paths to determine the access control quality and is equal to (1- w_u)

So *access control quality* is perfect if and only if *unknown path access control quality* is perfect and *known path access control quality* is perfect.

We further define

$$\textit{unknown path access control quality} = \text{effect of access paths on ACQ}\left(\frac{\textit{Insider access paths unknown to org}}{\textit{max reasonable paths}}\right)$$

$$\textit{known path access control quality} = \text{effect of access paths on ACQ}\left(\frac{\textit{extraneous access paths known to org}}{\textit{max reasonable paths}}\right)$$

where

- *max reasonable paths* is the number of access paths, known or unknown, beyond which no additional benefit is gained by the insider.
- *paths insider needs to do job* is the minimum number of access paths the insider needs to fulfill his job responsibilities

The function *effect of access paths on ACQ* is shown in Figure 18.

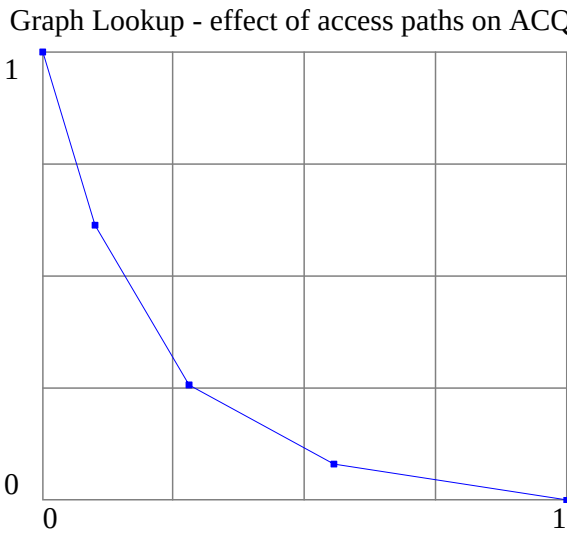


Figure 18: Function Defining Effect of Access Paths on Access Control Quality

The above assumes that an organization has perfect control of an employees' access if the following conditions hold:

1. The employees have no access paths unknown to the organization.
2. The employees have access only to access paths needed to do the job.

Employees with access to paths not meeting one of these two conditions indicate an access control lapse. The model's access control metric weighs access control lapses in condition 1 more heavily than those in condition 2. The graph of access

control over time is shown in Figure 19. This figure has the same general shape as given in the explanation for the iAssemble attack given in Figure 21. The rate of acting inappropriately online roughly captures the sharing of passwords and the installation of a backdoor prior to the termination and the unauthorized access and planting of the logic bomb during and after termination.

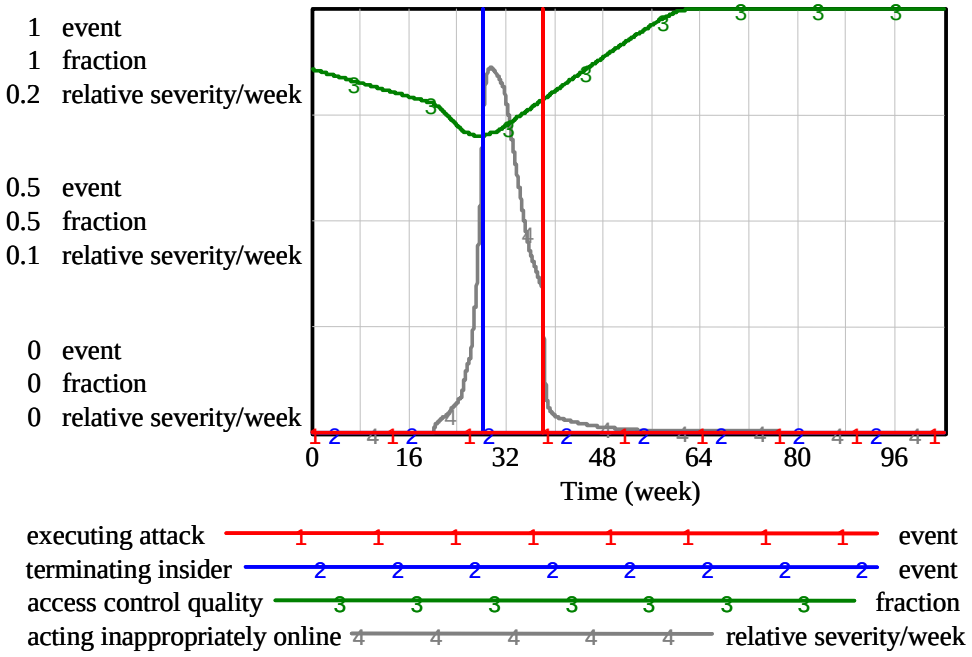


Figure 19: Explanation for iAssemble Attack (Simulation)

11 Conclusion

The MERIT project was initiated as a proof of concept—to determine whether or not an effective interactive learning environment could be developed to teach executives, managers, technical staff, human resource staff, and security officers the complex dynamics of the insider threat problem. An appropriate ILE must be intuitive enough to be easily understood by practitioners who have most likely never heard of system dynamics.

The steps required to develop the ILE are

- Gather and analyze extensive insider threat cases (completed—Insider Threat Study).
- Scope the problem for the model (completed—IT sabotage cases).
- Put together team of experts (completed—team consists of experts in insider threat, system dynamics, technical security, psychology).
- Build the model (in progress—current model described in this paper).
- Run simulations for initial testing and calibration of the model (in progress—some simulations described in this paper).
- Create training materials to accompany the model and ILE (student led development of training materials document as part of student project report [Desai 2006]).

At this point, the MERIT team feels confident that an effective model that conveys important lessons regarding insider threat has been created. The simulations accurately mimic the patterns and trends in the majority of the cases in the Insider Threat Study. Further calibration and validation of the model is still necessary before it can be released for educational or training use. In addition, extensive user interface testing will be required to develop an intuitive interface and accompanying training materials before the model can be used in an actual training class.

In addition to training, the MERIT team plans to present the model to experts in technical security, human resources, and organizational dynamics to calibrate it accurately so that it can be used for additional insights into the insider threat problem and effective countermeasures.

Appendix A: The iAssemble Training Case[11]

iAssemble sells computer systems directly to customers, building each system made-to-order and offering competitive prices. iAssemble has been doing extremely well and conducted an initial public offering (IPO) in 2001, after which its stock doubled.

ORGANIZATION BACKGROUND

iAssemble is headed by Chris Eagles, who is the Chief Executive Officer (CEO). Eagles started the company in 1997 with two of his friends, Carl Freeman and Caroline Thompson, who are now the Chief Financial Officer (CFO) and Chief Technical Officer (CTO), respectively. The company had continually hired experienced managers and employees over time.

Ian Archer, the malicious insider, was among the few employees who had been with iAssemble since its establishment. Archer started out as a computer specialist and technical assistant to the three original founders, Eagles, Freeman, and Thompson. When hired, Archer held certifications in personal computer (PC) hardware maintenance and operating system administration but did not possess a four-year, baccalaureate degree. He compensated for his lack of education with hard work and over the next four years he became the sole system administrator at iAssemble.

iAssemble grew at a moderate rate. Recognizing the need for qualified personnel, Eagles and Thompson began to hire experienced system administrators who could also function as project managers. Lance Anderson was hired as lead system administrator because of his education and qualifications, and James Allen was hired as a Junior System Administrator to share Archer's growing systems administration workload and responsibilities.

INSIDER SITUATION

Ian Archer had always been responsible for the software that ran the assembly machinery, and played an important role when iAssemble automated its PC assembly processes.

Archer's disgruntlement grew steadily over the course of several months due to the growth and associated changes at iAssemble. When Anderson was handpicked for the Lead Administrator position by Caroline Thompson, Archer began to feel confined in his current role and saw limited opportunities for advancement. Policy changes by Anderson meant that Archer could no longer work with the freedom he had always enjoyed. He began receiving detailed instructions on how to work and

[11] The iAssemble organization and case example are completely fictional; any resemblance to a real organization or insider threat case is unintentional.

felt “micro-managed.” Archer’s performance slumped, prompting Anderson and senior management to look to others, such as Allen, for important projects. As a result, Archer felt detached from the new culture at iAssemble, its leadership, and its continuing success.

Archer was assigned to mentor Allen and ensure his smooth assimilation within iAssemble’s culture. Archer and Allen worked on a few small projects together but Archer felt that the projects were too menial for his technical skills. He found Allen to be nearly as technically competent as himself, which contributed to his frustration. While working on one of these projects, Allen shared the password to his personal desktop machine at iAssemble, named *Kilimanjaro*, with Archer. Sharing the password enabled each of them to access the project files when the other was out of the office.

Archer’s disgruntlement grew, and he openly proclaimed that Anderson was just a figurehead. If anyone disagreed he verbally abused them until they backed down and apologized. He even bottlenecked projects on purpose on several occasions, stalling his work on the project to ensure Anderson and the project team missed project milestones. Archer received a written warning from Thompson after several co-workers formally complained. Enraged by this, he had a heated argument with a team member who then quit the very next day, citing Archer as the reason for his resignation. Archer was suspended for a day without pay and received a cut in his salary.

At this point Thompson became more cautious regarding Archer and wanted to fire him. Anderson warned that firing a disgruntled system administrator was a complicated task. Almost every company he worked for had access control gaps that would allow an ill-tempered, ex-employee to cause system damage. He suspected such a scenario existed for iAssemble, and in the face of Archer’s firing, could be risky. Anderson believed that yearly audits iAssemble conducted lacked proper vigor and documentation, and there was no way to be certain that they had reduced their risk for sabotage by a former employee. A decision was made to increase audits of access control quality and access management. The audits would begin immediately.

INSIDER ATTACK

After the blowup with his team member and the subsequent salary cut, Archer had the feeling that he would soon be fired. He decided that he needed to have the means to get back at iAssemble in the event that his worst fears came true.

The audits revealed a great deal about iAssemble’s access management. Many access paths, both of present and past employees, were discovered which should have been disabled. Dummy accounts were discovered which were created for testing and debugging purposes but never deleted; a few had even been created by Archer. These access paths were promptly disabled. Thompson felt there was steady progress being made and deemed the audit an excellent decision.

In the meantime, Archer planted a backdoor on the main machinery server that provided him with unauthorized access. Archer's immediate anger with iAssemble was somewhat alleviated after this act, but he was still disgruntled.

Archer became infuriated when he overheard that management was planning to fire him. He decided to wait until after his termination, and then seek revenge. Two days later, on December 14, 2001, Archer was fired and his access disabled. Unfortunately, iAssemble managers were not aware that he knew the password to *Kilimanjaro*, James Allen's iAssemble machine, and did not think to change it. Archer went home the night he was fired and successfully logged into *Kilimanjaro*. He then used his backdoor on the machinery server to plant the logic bomb. He did the same on both of the backup servers. He cleverly set it to go off two months from that date to deflect suspicion from him. Figure 20, below, shows Ian Archer's attack method.

On February 14, 2002, the logic bomb went off, deleting all files on the machinery server and its backup servers, leaving the assembly lines at iAssemble frozen.

ORGANIZATION'S RESPONSE

When investigators suggested the possibility of insider attack, management was puzzled as to how that was possible in light of the increased monitoring, policies and best practices in place at iAssemble. Eventually, the system logs were used to trace the access to the machinery server from *Kilimanjaro*. James Allen claimed that he was innocent and explained that he had given the password to his machine to Archer when they worked together.

Ian Archer was soon arrested, but iAssemble was left on shaky ground. Their share prices plummeted and their image in the market was blemished.

Management concluded that growth had taken its toll at iAssemble with the company facing access control and employee disgruntlement issues. With a whopping growth in sales figures of 68% over the past quarter, iAssemble was rapidly hiring good and efficient people. However, this situation created competition for positions within iAssemble, leading to job dissatisfaction among some employees. Access control quality also eroded over time with employees under pressure to meet deadlines and subsequently violating security policies. Figure 21 depicts the dynamics that occurred at iAssemble.

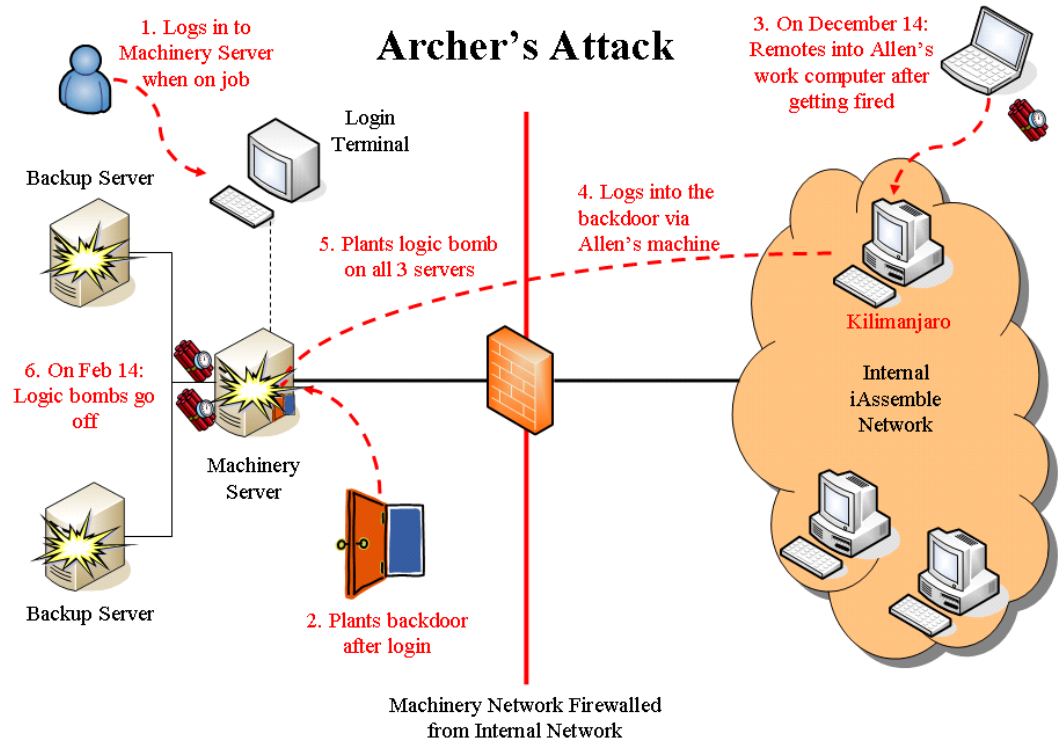


Figure 20: Insider's Method of Attack

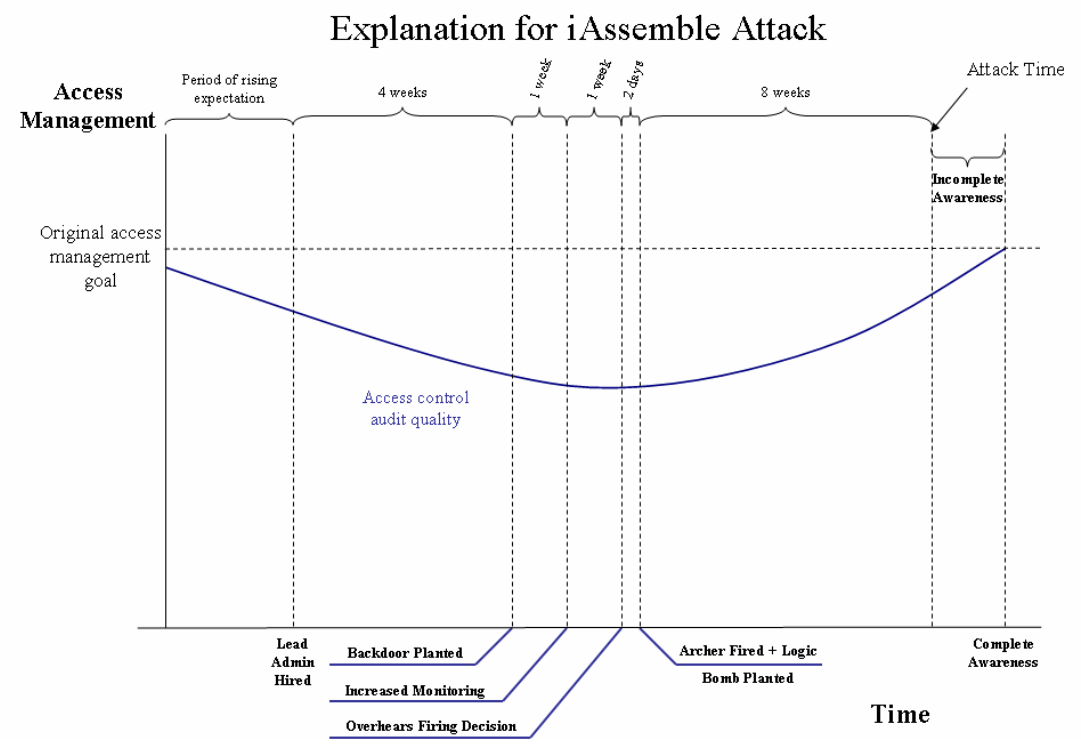


Figure 21: Explanation for iAssemble Attack

Despite iAssemble's efforts to maintain information security best practices, security policy enforcement became lax in support of the culture of growth at all costs. Hence, access control deteriorated over time.

When iAssemble managers realized that they needed to increase monitoring and audits, they started out on the right foot. The steps they took ensured that management's knowledge of their employees' access paths increased. However, the skew between management's knowledge of Archer's access paths and the actual access paths that Archer had, was not fully overcome by the time Archer was fired, and iAssemble could not disable all of his access paths in time. Hence, Archer was able take advantage of the residual access that he had, as shown in Figure 22.

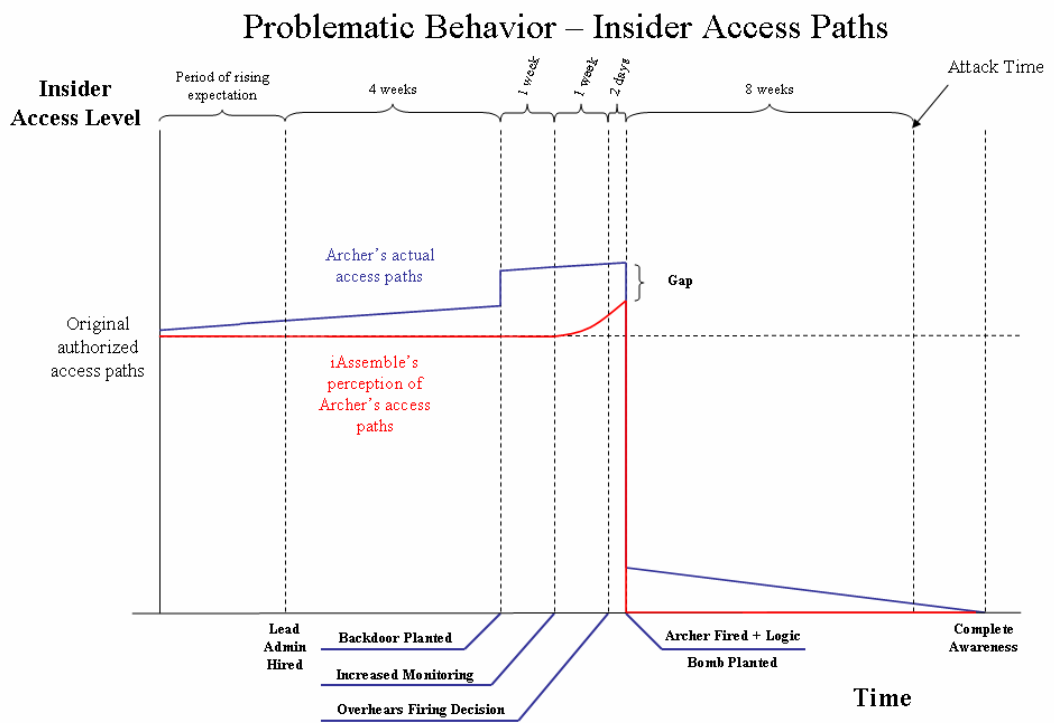
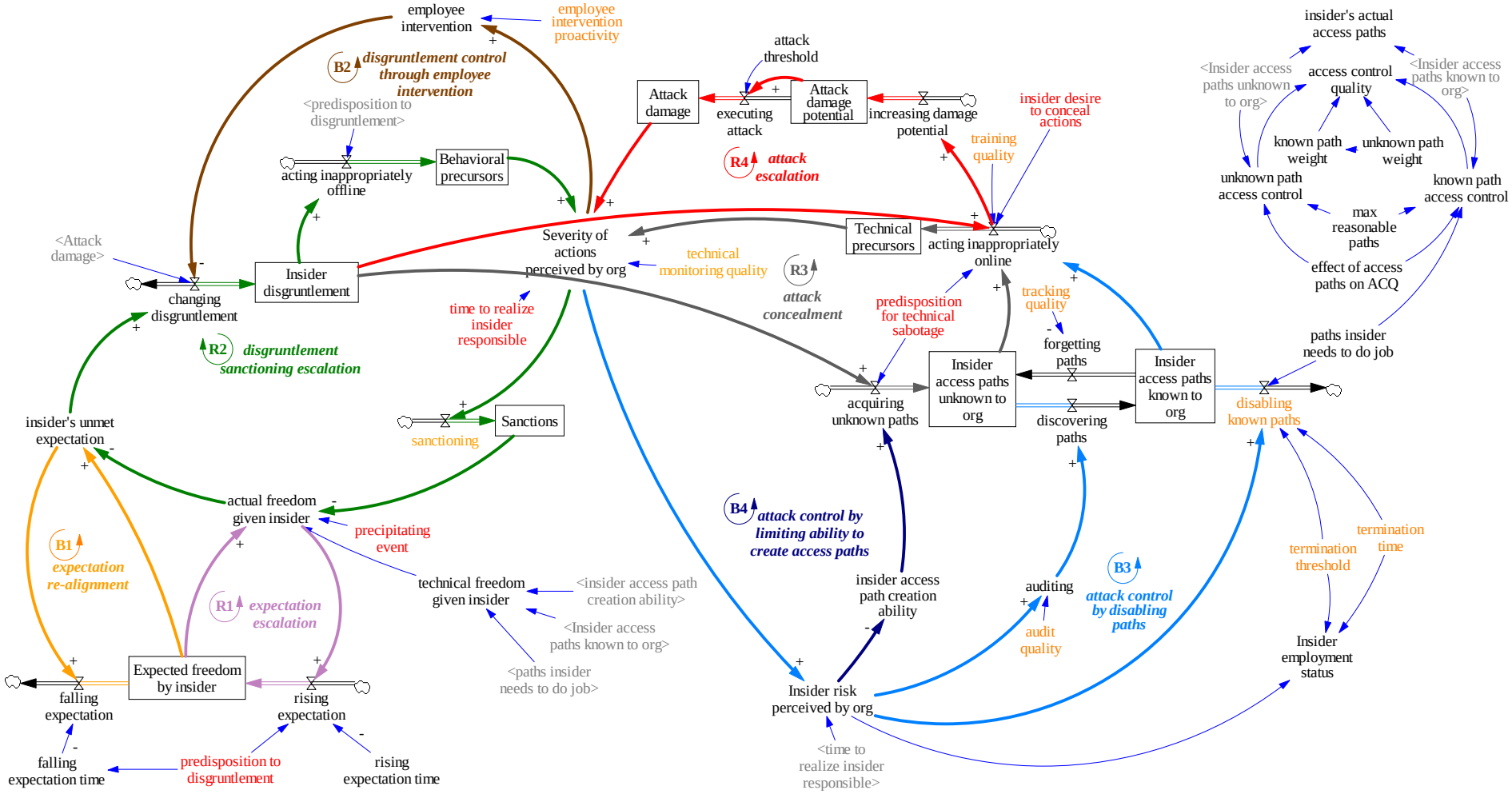


Figure 22: Analysis of Insider Access Level

Appendix B: Simulation Model Overview



References

[Argyris 1974]
Argyris, C., & Schon, D. *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey Bass, 1974.

[Anderson 2004]
Anderson, D. F.; Cappelli, D. M.; Gonzalez, J. J.; Mojtahedzadeh, M.; Moore, A. P.; Rich, E., Sarriegui, J. M.; Shimeall, T.J.; Stanton, J. M.; Weaver, E.; & Zagonel, A. "Preliminary System Dynamics Maps of the Insider Cyber-Threat Problem." *Proceedings of the 22nd International Conference of the System Dynamics Society*, July 2004.
http://www.cert.org/archive/pdf/InsiderThreatSystemDynamics.pdf.

[Band 2006]
Band, S.R.; Capelli, D.M.; Fischer, L.F.; Moore, A.P.; Shaw, E.D.; & Randall, F. T. *Comparing Insider IT Sabotage and Espionage: A Model-Based Analysis* (CMU/SEI-2006-TR-026). Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 2006.
http://www.sei.cmu.edu/publications/documents/06.reports/06tr026.html.

[Desai 2006]
Desai, A.G. "Insider Threat Dynamics." *Project Report for Master of Science in Information Networking.* Pittsburgh, PA*:* Information Networking Institute, Carnegie Mellon University, 2006.

[Huber 2002]
Huber, C.; Cooke, F.; Smith, J.; Paich, M.; Pudar, N.; & Barabba, V. "A Multimethod Approach for Creating New Business Models: The General Motors OnStar Project," *Interfaces 32*, 1: 20-34, 2002.

[Groessler 2004]
Groessler, A. "Don't let history repeat itself – methodological issues concerning the use of simulators in teaching and experimentation." *System Dynamics Review 20*, 3: 263-274, 2004.

[Janis 1977]
Janis, I. & Mann, L. *Decision Making: A Psychological Analysis of Conflict, Choice and Commitment*. New York, NY: The Free Press, 1977.

[Kahneman 1982]
Kahneman, D.; Slovic, P.; & Tversky, A. *Judgment Under Uncertainty: Heuristics & Biases*. Cambridge, UK: Cambridge University Press, 1982.

[Keeney 2005]
Keeney, M. M.; Kowalski, E. F.; Cappelli, D. M.; Moore, A. P.; Shimeall, T. J.; & Rogers, S. N. "Insider Threat Study: Computer System Sabotage in Critical Infrastructure Sectors." *Joint SEI and U.S. Secret Service Report*, 2005.
http://www.cert.org/archive/pdf/insidercross051105.pdf.

[Lane 1995]
Lane, D. "On a resurgence of management simulations and games." *The Journal of the Operational Research Society, 46*, 5: 604-625, 1995.

[Meadows 1974]
Meadows, D. L.; Behrens, W. W.; Meadows D. H.; Naill, R. F.; Randers, J.; & Zahn, E.K.O. *Dynamics of Growth in a Finite World*. Cambridge, MA: Wright-Allen Press, Inc., 1974.

[Melara 2003]
Melara, C.; Sarriegui, J.M.; Gonzalez, J.J.; Sawicka, A.; & Cooke, D.L. "A system dynamics model of an insider attack on an information system." *Proceedings of the 21st International Conference of the System Dynamics Society*. New York, NY, July 20-24, 2003. Albany New York: Systems Dynamic Society, 2003.

[Moore 2005]
Moore, A.P. & Cappelli, D.M. "Analyzing Organizational Cyber Threat Dynamics." *Proceedings of the Workshop on System Dynamics of Physical and Social Systems for National Security*. Chantilly, VA, April 21-22, 2005.

[Randazzo 2004]
Randazzo, M. R.; Keeney, M.M.; Kowalski, E. F.; Cappelli, D.M.; & Moore, A.P. "Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector" *Joint SEI and U.S. Secret Service Report*, 2004. http://www.cert.org/archive/pdf/bankfin040820.pdf.

[Rich 2005]
Rich, E.; Martinez-Moyano, I. J.; Conrad, S.; Cappelli, D.M.; Moore, A.P.; Shimeall, T.J.; Andersen, D.F.; Gonzalez, J. J.; Ellison, R.J.; Lipson, H.F.; Mundie, D.A.; Sarriegui, J. M.; Sawicka, A.; Stewart, T. R.; Torres, J. M.; Weaver, E. A.; & Wiik, J. "Simulating Insider Cyber-Threat Risks: A Model-Based Case and a Case-Based Model." *Proceedings of the 23rd International Conference of the System Dynamics Society*, Boston, MA, July 2005. Albany, NY: Systems Dynamics Society, 2005.

[Rosenthal 1992]
Rosenthal, R. & Jacobson, L. *Pygmalion in the classroom* (expanded edition). New York, NY: Irvington, 1992.

[Sterman 2000]
Sterman, John D. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York, NY: McGraw-Hill, 2000.

[Sterman 2006]
Sterman, J.D. "Learning from evidence in a complex world." *American Journal of Public Health, 96*, 3: 505-514, 2006.

[Tedeschi 1981]
Tedeschi, J. T. *Impression management theory and social psychological research.* New York, NY: Academic Press, 1981.

[Wason 1966]
Wason, P. C. "Reasoning." In B. M. Foss (Ed.) *New Horizons in Psychology*. B. M. Foss (ed.). Hammondsworth, UK: Penguin, 1966.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	March 2007	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Management and Education of the Risk of Insider Threat (MERIT): Mitigating the Risk of Sabotage to Employers' Information, Systems, or Networks	FA8721-05-C-0003

6. AUTHOR(S)

Dawn M. Cappelli; Akash G. Desai; Andrew P. Moore; Timothy J. Shimeall; Elise A. Weaver; & Bradford J. Willke

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2006-TN-041

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
HQ ESC/XPK 5 Eglin Street Hanscom AFB, MA 01731-2116	

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

The Insider Threat Study, conducted by the U.S. Secret Service and Carnegie Mellon University's Software Engineering Institute CERT Program, analyzed insider cyber crimes across U.S. critical infrastructure sectors. The study indicates that management decisions related to organizational and employee performance sometimes yield unintended consequences that increase risk of insider attack. The problem is exacerbated by a lack of tools for understanding insider threat, analyzing risk mitigation alternatives, and communicating results. To develop such tools is the goal of Carnegie Mellon University's Management and Education of the Risk of Insider Threat (MERIT) project. MERIT uses system dynamics to model and analyze insider threats and produce interactive learning environments. These tools can be used by policy makers, security officers, information technology and human resource personnel, and management. The tools help these users to understand the problem and assess risk from insiders based on simulations of policies, and on cultural, technical, and procedural factors. This technical note describes the MERIT insider threat model and simulation results.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Insider Threat Study, insider attack, MERIT, simulation-based learning, access control	61

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. Z39-18 298-102